A SIMPLE INNOVATIVE METHOD OF INTERPRETING THE BREAK-THROUGH CURVES IN INSTANTANEOUS SOURCE COLUMN TESTS

A Thesis

by

KAIYI ZHANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Hongbin Zhan
Committee Members,	Peter Knappett
	David Sparks
Head of Department,	John R. (Rick) Giardino

May 2019

Major Subject: Water Management and Hydrological Science

ABSTRACT

Instantaneous injection tracer tests can be used effectively to determine solute transport parameters in porous media such as pore velocities and dispersivities, which are usually estimated with curve-fitting methods. This study proposes a simple method to estimate conservative and reactive solute transport parameters in one-, two- and three- dimensional domains with uniform flow fields based on combination of certain selected observation times. This method requires fewer measured data than traditional curve-fitting methods. The accuracy of the method in one-dimensional domain depends on the selection of three time points, which is a key factor for the proposed method of this study. Based on the uncertainty analysis the proposed method appears to be a robust and creditable assessment tool applicable for estimating parameters with acceptable estimation errors.

The proposed method is applied on laboratory sand column tests. The error of dispersivity between the proposed method and the curve-fitting method would be less if the velocity in the column is lower. For velocity of 0.10cm/min, the error of dispersivity is 18%, while with a lower velocity of 0.05cm/min, the error of dispersivity would be 10% less (8%). The results indicate that the estimated pore velocities and dispersivities are almost the same to their counterparts of the curves-fitting method. This method can be employed easily by scientists and practitioners for parameter estimations in laboratory column experiments if advection-dispersion equation is applicable. Limitations of the study have also been addressed.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Hongbin Zhan, and my committee members, Dr. David Sparks, and Dr. Peter Knappett, for their guidance and support throughout the course of this research.

I would like to thank Dr. Guiming Dong for his support in laboratory injecting tracer tests.

Thanks also to my friends and colleagues who helped on my research and their support which are so important to my study. Thanks to Qiong Su, who is now pursuing her Ph.D. degree in Water Management and Hydrological Science, I sincerely admire her encouragement and support with humility, faithful words with respect. Finally, thanks to my mother and father for their encouragement.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Hongbin Zhan, Dr. Peter Knappett, and Dr. David Sparks of the Department of Geology & Geophysics, Texas A&M University, and Dr. Guiming Dong of School of Resources and Geosciences, China University of Mining and Technology.

All other work conducted for the thesis was completed by the student.

Funding Sources

The laboratory tracer tests included was partially supported with the research grants from the program "The Social Development-Science & Technology Demonstration Projects" sponsored by Department of Science and Technology of Jiangsu Province (BE2015708).

Graduate study was partly supported by academic scholarships awarded by the Department of Water Management and Hydrological Science, and China Scholarship Council.

TABLE OF CONTENTS

Page

NOMENCLATURE

ADE	Advection-dispersion equation.
BTCs	Breakthrough curves.
CFM	Curve-fitting method.
C	Concentration of the adsorbate in solution [$M \cdot L^{-3}$].
u	Average pore velocity [$L \cdot T^{-1}$].
q	Darcy's velocity [$L \cdot T^{-1}$].
n	Effective porosity [dimensionless].
x	Longitudinal dimension along the direction of flow [L].
D	Dispersion coefficient [$L^2 \cdot T^{-1}$].
D_L	Longitudinal (x-direction for this study) dispersion coefficient [L^2/T].
α_L	Longitudinal dispersivity [L].
D_0	Effective molecular diffusion coefficient in porous media, which is usually computed as a product of the free-water molecular diffusion coefficient and a tortuosity coefficient between 0 and 1.
M	Total amount of solute contained in the slug [M].
A	Area of the injection cross-section that is perpendicular to the flow direction [L^2].
L	Length of injection in the direction perpendicular to the xy plane [L].

D_T	Transverse (y-direction for this study) dispersion coefficient [L^2/T].
α_T	Transverse dispersivity [L].
D_{HT}	Horizontally transverse (y-direction for this study) dispersion coefficient [L^2/T].
α_{HT}	Horizontally transverse dispersivity [L].
D_{VT}	Vertically transverse (z-direction for this study) dispersion coefficient [L^2/T].
α_{VT}	Vertically transverse dispersivity [L].
R	Constant retardation factor [−].
ρ_d	Bulk density of porous media [M/L^3].
K_d	Distribution coefficient [L^3/M].
t_i	Observation time i [T].
c_i	Concentration of tracer at observed time t_i [M/L^3].
t_m	Peak time [T].
c_m	Peak concentration of tracer at peak time t_m [M/L^3].
N	Number of pairs of data of observation time and corresponding concentration [-].

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1 Background

The contaminant transport in porous media is commonly assumed to be governed by the advection-dispersion equation (ADE) (Bear, 1972), which is one of many conceptual models for interpreting transport phenomena. ADE has been widely applied in many laboratory and field studies, and can be an applicable model to describe the solute transport in many cases. The simplest form of ADE for one-dimensional (1-D) transport with a linear sorption (or a constant retardation factor) without sink/source can be expressed as:

$$R\frac{\partial y}{\partial x} = \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) - u\frac{\partial C}{\partial x} \tag{1}$$

where C is the concentration of the adsorbate in solution [M·L^{-3}], u=q/n is the average pore velocity [L·T^{-1}], q is Darcy's velocity [L·T^{-1}], n is effective porosity, t is time [T], x is the longitudinal dimension along the direction of flow [L], D is the dispersion coefficient [L^2·T^{-1}], R is a retardation factor associated with the linear sorption [dimensionless].

Based on this theory, instantaneous tracer injection test is used widely to determinate contaminant transport parameters in porous media or subsurface environments (Bear, 1961; Mackay et al., 1986). The transport parameters including porosities, pore velocities, and dispersivities are very important to investigate the fate and transport of the contaminants and colloid in the subsurface (Ma et al., 2018; Boy-Roura et al., 2018; Han et al., 2017; Wu et al., 2017; Lv et al.,2016; Weaver et al., 2016), which are usually estimated with a proper curve-fitting method (CFM). Liang

et al. (2018) proposed a simple method to estimate conservative and reactive solute transport parameter in one-, two-, and three- dimensional domains with uniform flow fields based on peak times of slug tracer tests, and this method is hereinafter called the Liang's method.

The Liang's method requires fewer measured data than traditional CFMs, and accuracy of the method depends on the time-interval of measurement. In recent ten years ADE has been criticized by advocates of nonlocal theories for not well interpreting the observed solute transport in some field sites. For instance, Neuman and Tartakovsky (2009) focused on flow processes in heterogeneous media and stated that the estimated transport parameters from fitting ADE to the measured BTCs often become spatially and temporally dependent and that the observed BTCs exhibit tailing that is too strong for the ADE. Nevertheless, those nonlocal theories providing an alternative interpretation for solute transport in the heterogeneous media are still difficult to solve the real-world applications due to the complicated mathematic model, especially for the three-dimensional (3-D) problems. Ambiguity of physics behind some parameters of these nonlocal models has still not been fully resolved for applications. ADE is commonly used for dealing with many practical transport problems because of its simplicity, albeit it is found to be problematic in some cases, particularly for transport in highly heterogeneous porous media.

Analytical solutions of ADE for instantaneous tracer injection tests with uniform flow fields in one- (1-D), two- (2-D) and three dimensional (3-D) domains were derived by many previous studies (De Josselin De Jong, 1958; Domenico and Schwartz, 1990; Sauty, 1980). The solute transport parameters are usually estimated

using a CFM that fits the observed tracer concentrations at different times at a downstream position (the breakthrough curves or BTCs) with an appropriate theoretical (analytical) solution. The curve-fitting procedure is a common practice for parameter estimation, it also has several disadvantage: Firstly, it requires many measured concentrations to capture the entire BTCs, leading to a high cost of data collection, especially for the 2-D and 3-D cases. Secondly, it involves an optimized operation that sometimes falls into the local optimal estimates rather than the global ones.

1.2 Problem statement

1-D case: In a homogenous porous medium which is free of solutes initially, the groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$ and will not affect the conservative tracer test results. The analytical solution of BCTs for instantaneous source injected at $x = 0$ can be written as follows (Sauty, 1980)

$$c(x,t) = \frac{M}{2An\sqrt{D_L \pi t}} \exp\left[-\frac{(x-ut)^2}{4D_L t}\right] \quad (2)$$

Initial and boundary conditions used for above solution are listed below:

$$c(x,0) = 0 \quad x \geq 0 \quad (3)$$

$$c(0,t) = 0 \quad t \geq 0 \quad (4)$$

$$c(\infty,t) = 0 \quad t \geq 0 \quad (5)$$

where c is the solute concentration [M/L^3]; $D_L = \alpha_L |u| + D_0$ is the longitudinal (x-direction for this study) dispersion coefficient [L^2/T]; α_L is the longitudinal dispersivity [L]; u is the average groundwater flow velocity (or the pore velocity) and is along the x-axis [L/T]; D_0 is the effective molecular diffusion coefficient in porous

media, which is usually computed as a product of the free-water molecular diffusion coefficient and a tortuosity coefficient between 0 and 1; M is the total amount of solute contained in the slug; A is the area of the injection cross-section that is perpendicular to the flow direction [L^2]; n is the effective porosity [−].

2-D case: In a homogenous porous medium which is free of solutes initially, the groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$ and $y = \pm\infty$, and will not affect the tracer test results. The analytical solution of BTCs for conservative instantaneous source injected at $x = 0$ and $y = 0$ can be written as follows (De Josselin De Jong, 1958; Sauty, 1980)

$$c(x, y, t) = \frac{M}{4\pi Lnt\sqrt{D_L D_T}} \exp\left[-\frac{(x-ut)^2}{4D_L t}\right] \exp\left[-\frac{y^2}{4D_T t}\right] \quad (6)$$

where $D_T = \alpha_T|u| + D_0$ is the transverse (y-direction for this study) dispersion coefficient [L^2/T]; α_T is the transverse dispersivity [L]; L is the length of injection in the direction perpendicular to the xy plane [L], and L equals the aquifer thickness for a fully penetrating vertical line source. Eq. (6) describes tracer transport in a 2-D space with advection and dispersion in the x-direction and dispersion in the y-direction as well.

3-D case: Groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$, $y = \pm\infty$, and $z = \pm\infty$, and will not affect the tracer test results. The analytical solution of BCTs for conservative instantaneous source injected at $x = 0$, $y = 0$ and $z = 0$ can be written as follows (Park and Zhan, 2001)

$$C(x, y, z, t) = \frac{M}{(4\pi t)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}} \exp\left[-\frac{(x-ut)^2}{4D_L t}\right] \exp\left(-\frac{y^2}{4D_{HT} t}\right) \exp\left(-\frac{z^2}{4D_{VT} t}\right) \quad (7)$$

where $D_{HT} = \alpha_{HT}|u| + D_0$ is the horizontally transverse (y-direction for this study) dispersion coefficient [L^2/T]; α_{HT} is the horizontally transverse dispersivity [L]; $D_{VT} = \alpha_{VT}|u| + D_0$ is the vertically transverse (z-direction for this study) dispersion coefficient [L^2/T]; α_{VT} is the vertically transverse dispersivity [L]. Eq. (7) describes tracer transport in a 3-D space with advection and dispersion in the x-direction and dispersion in both the y-direction and the z-direction.

For reactive solutes, here we consider a linear sorption isotherm and a first-order decay for the tracer transport in porous media with uniform flow fields. Analytical solutions for the 1-D, 2-D, and 3-D cases can be written as follows respectively,

$$c(x,t) = \frac{M}{2AnR\sqrt{D_L \pi t/R}} \exp\left[-\frac{(x-ut/R)^2}{4D_L t/R} - \lambda t\right] \tag{8}$$

$$c(x,y,t) = \frac{M}{4\pi Lnt\sqrt{D_L D_T}} \exp\left[-\frac{(x-ut/R)^2}{4D_L t/R} - \lambda t\right] \exp\left(-\frac{y^2}{4D_T t/R}\right) \tag{9}$$

$$c(x,y,z,t) = \frac{M}{(4\pi t)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}} \exp\left[-\frac{\left(x-\frac{ut}{R}\right)^2}{\frac{4D_L t}{R}} - \lambda t\right] \exp\left(-\frac{y^2}{\frac{4D_{HT} t}{R}}\right) \exp\left(-\frac{z^2}{\frac{4D_{VT} t}{R}}\right) \tag{10}$$

where $R = 1 + \rho_d K_d/n$ is the constant retardation factor [−];ρ_d is the bulk density of porous media [M/L^3]; K_d is the distribution coefficient [L^3/M]; λ is the rate constant for the first-order decay in the liquid [T^{-1}] and it is usually given in terms of the half-time as $\lambda = (\ln 2)/t_{1/2}$, where $t_{1/2}$ is the half-life of a radioactive or a biodegradable tracer [T].

The above analytical solutions of ADE show that the concentrations are normally distributed functions of spatial coordinates at any given time. However, the relationship between solute concentration and time at a certain observed point does not follow the normal-distribution pattern. This implies that the conventional use of

breakthrough curves (BTCs), which are concentration-time curves at a given observed point, will exhibit non-normal distribution pattern. The degree of deviation of BTCS from the normal distribution pattern depends mostly on the Peclet number which is a dimensionless number reflecting the relative importance of advective transport versus dispersive transport. The non-normality of the BTCs offers some challenges and opportunities for parameter interpretation. The purpose of this thesis is to seek a simple general method for interpreting the BTCs by acknowledging the non-normality of the BTCs.

1.3 Motivation

The curve-fitting method is usually used to estimate the solute transport parameters by fitting the observed BTCs with an appropriate analytical solution. To achieve this purpose, some parameter estimation methods have been programmed such as CXTFIT (Parker and Vangenuchten, 1984; Toride et al., 1995), CXTANNEAL (Li et al., 1999), UCODE (Poeter and Hill, 1999), and PEST (Doherty, 2001). A FORTRAN IV computer program CXTFIT was developed firstly in CXTFIT 1.0 code by Parker and Vangenuchten (1984) to estimate solute transport parameters using a nonlinear lest squares parameter optimization method. The program can be used to solve the inverse problem by fitting the analytical solution of ADE with the BTCs under 1-D steady-state flow condition. Toride et al. (1995) improved the CXTFIT 1.0 code and proposed a CXTFIT 2.0 code that was used to solve the inverse problem by a nonlinear least squares inversion method according to Levenberg-Marquardt technique, which is also known as the damped least-squares method. The CXTFIT 2.0 code includes a greater number of analytical solutions to various initial and boundary

conditions then later it was further improved by Li et al. (1999) who developed CXTANNEAL. CXTANNEAL is a program for analyzing contaminant transport in soils. The code, written in Fortran 77, is a modified version of CXTFIT, and is a commonly used package for estimating solute transport parameters in soils. The improvement is that it includes simulated annealing as the optimization technique for curve fitting. Tests with hypothetical data show that CXTANNEAL performs better than the original code in searching for optimal parameter estimates. To reduce the computational time, a parallel version of CXTANNEAL (CXTANNEAL_P) was also developed. US Geological Survey computer program UCODE (Poeter and Hill, 1999) was developed to perform inverse modeling as a parameter-estimation problem, using nonlinear regression. Estimated parameters can be defined flexibly with user-specified functions, and the nonlinear regression problem is solved by minimizing a weighted least-squares objective function with respect to the parameter values using a modified Gauss–Newton method. Both UCODE (Poeter and Hill, 1999) and PEST (Doherty, 2001) are automatic calibration procedure using weighted nonlinear regression by minimizing an objective function related to the square difference between observed and simulated variables. Moreover, UCODE and PEST algorithms have been widely applied on the calibration of hydrologic models.

Although the above-mentioned curve-fitting procedure is a common practice for parameter estimation, it has several disadvantages which cannot be neglected and these disadvantages could be summarized as follows. Firstly, it requires scientists and practitioners to measure concentrations for the purpose to capture the entire BTCs, and such whole data collection process is expensive, especially for the 2-D and 3-D cases,

leading to a high cost. Secondly, it involves an optimized operation that sometimes falls into the local optimal estimates rather than the global ones (Liang et al. 2018). Both UCODE and PEST methods are local search unconstrained calibration methods, the underlying drawback that the resulting parameter values can be located in a local minimum of the objective function will lead to obtaining physically unrealistic parameter values from the perspective of global scale goal. So if scientists and practitioners want to apply the above-mentioned curve-fitting methods, more efficient sampling techniques or powerful and robust global optimization schemes should be considered and incorporated.

Liang et al. (2018) considered the main disadvantages of normal-used curve-fitting methods, analyzed the derivation result of the concentration over time based on the analytical solutions of ADE, and proposed a simple method (Liang et al. 2018) depending on the peak time and corresponding peak concentration on the BTCs to estimate parameters by operating the closed-form algebraic functions. He summarized and stated that one needs to determine two observation points and obtain the pair of the peak time and peak concentration on each observation point at first, later one could substitute these data into his formulas to estimate parameters in 1-D domain. In a word, if one wants to apply his method into practice, both conservative and reactive instantaneous solutes need at least two, three, and four observed points for the 1-D, 2-D, and 3-D case, respectively. Nevertheless, the accuracy of his parameter estimation method largely relies on the accurate determination of the peak time, which is closely related to the time-interval of measurement. In practice, the time-interval of measurement is always a finite value that will inevitably induce either underestimation

or overestimation of the actual peak time. In this thesis, I propose a simple and more straightforward method to estimate the transport parameters including pore velocities and dispersivities for conservative and reactive solutes in 1-D, 2-D and 3-D domains with uniform flow fields based on the measured concentrations of the instantaneous tracer tests. This method requires fewer measured data than the traditional curve fitting methods, and neither optimized operations nor the pairs of data exactly at the peak time and corresponding peak concentration are required. This method can be employed straightforwardly by scientists and practitioners for parameter estimations in laboratory column tests provided that ADE is applicable.

1.4 Objectives

In this thesis, I plan to conduct a combined analytical investigations, synthetic data and laboratory instantaneous source column tests to achieve the following objectives:

Objective 1. I will propose a simple innovative method to interpret BTCs for instantaneous source solute transport, and demonstrate the robustness of the method for parameter estimations for both conservative and reactive solutes in 1-D, 2-D, and 3-D cases, respectively.

Objective 2. I will apply the proposed method to laboratory instantaneous source column tests by developing MATLAB scripts to search for the best sampling time particularly for conservative solute in 1-D case. Based on the laboratory data, I will develop MATLAB scripts for both CFM and the Liang's method for parameter estimation, then compare these two methods with the proposed one, and conduct an uncertainty analysis.

Objective 3. I will also check and illustrate how some associated factors would affect the shape and skewness of BTCs, specifically the average pore velocity u and dispersivity α for conservative solute transport. Based on the laboratory column tests, I will demonstrate how these factors would affect the time sets selection for parameter estimation purpose.

2. METHODOLOGY

2.1 Conservative Solutes

1-D case: In a homogenous porous medium which is free of solutes initially, the groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$ and will not affect the conservative tracer test results. The analytical solution of BCTs for instantaneous source injected at $x = 0$ can be written as Eq. (2). If input mass M is known, there are three parameters (u, n, α_L) that need to be estimated. Therefore, we could use minimum three concentrations selected at three different times, which are (c_1, t_1, x_1), (c_2, t_2, x_2), and (c_3, t_3, x_3) to figure out those three parameters (u, n, α_L). To do so, one can substitute them into Eq. (2):

$$c_1 = \frac{M}{2An\sqrt{D_L \pi t_1}} \exp\left[-\frac{(x_1 - ut_1)^2}{4D_L t_1}\right] \tag{11}$$

$$c_2 = \frac{M}{2An\sqrt{D_L \pi t_2}} \exp\left[-\frac{(x_2 - ut_2)^2}{4D_L t_2}\right] \tag{12}$$

$$c_3 = \frac{M}{2An\sqrt{D_L \pi t_3}} \exp\left[-\frac{(x_3 - ut_3)^2}{4D_L t_3}\right] \tag{13}$$

Dividing Eq. (11) by Eq. (12), and dividing Eq. (12) by Eq. (13) one can have

$$\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}} = \exp\left[\frac{(x_2 - ut_2)^2}{4D_L t_2} - \frac{(x_1 - ut_1)^2}{4D_L t_1}\right] \tag{14}$$

$$\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}} = \exp\left[\frac{(x_3 - ut_3)^2}{4D_L t_3} - \frac{(x_2 - ut_2)^2}{4D_L t_2}\right] \tag{15}$$

Eq. (14) and Eq. (15) could be transformed into Eq. (16) and Eq. (17) as follows respectively,

$$ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right) = \frac{(x_2 - ut_2)^2}{4D_L t_2} - \frac{(x_1 - ut_1)^2}{4D_L t_1} \tag{16}$$

$$ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right)=\frac{(x_3-ut_3)^2}{4D_Lt_3}-\frac{(x_2-ut_2)^2}{4D_Lt_2} \tag{17}$$

Dividing Eq. (16) by Eq. (17) one generates Eq. (18) as follow,

$$\frac{ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right)}{lln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right)}=\frac{\frac{(x_2-ut_2)^2}{t_2}-\frac{(x_1-ut_1)^2}{t_1}}{\frac{(x_3-ut_3)^2}{t_3}-\frac{(x_2-ut_2)^2}{t_2}} \tag{18}$$

Eq. (18) could be transformed into quadratic equation only with one unknown parameter u,

$$(\omega_1\beta_2-\omega_2\beta_1)u^2\text{-}(\varepsilon_1\beta_2-\varepsilon_2\beta_1)u+(\gamma_1\beta_2-\gamma_2\beta_1)=0 \tag{19}$$

where

$$\omega_1=t_2-t_1,\ \varepsilon_1=2(x_1-x_2)\ ,\ \gamma_1=\frac{x_2^2}{t_2}-\frac{x_1^2}{t_1}\ ,\ \beta_1=\ln(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}})\ ,$$

$$\omega_2=t_3-t_2,\ \varepsilon_2=2(x_2-x_3)\ ,\ \gamma_2=\frac{x_3^2}{t_3}-\frac{x_2^2}{t_2}\ ,\beta_2=\ln(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}).$$

We could solve Eq. (19) and calculate u:

$$u=\frac{(\varepsilon_2\beta_1-\varepsilon_1\beta_2)+\sqrt{(\varepsilon_1\beta_2-\varepsilon_2\beta_1)^2-4(\omega_1\beta_2-\omega_2\beta_1)(\gamma_1\beta_2-\gamma_2\beta_1)}}{2(\omega_1\beta_2-\omega_2\beta_1)} \tag{20}$$

Then we could substitute u into Eq. (16) for D_L

$$D_L=\frac{\omega_1u^2+\varepsilon_1u+\gamma_1}{4\beta_1} \tag{21}$$

Later substituting u and D_L into Eq. (11) one could have

$$n=\frac{M}{2Ac_1\sqrt{D_L\pi t_1}}\exp\left[-\frac{(x_1-ut_1)^2}{4D_Lt_1}\right] \tag{22}$$

This procedure is also applicable if spilled contaminant source mass *M* is unknown. To do so, we have to measure the effective porosity *n* instead of curve-fitting the *n* value as described above, then the source mass *M* could be estimated as Eq. (23)

$$M = 2Ac_1\, n\sqrt{D_L \pi t_1}\exp\left[\frac{(x_1-ut_1)^2}{4D_L t_1}\right] \quad (23)$$

A special case is that we choose the same observed point but different time, meaning that $x_1 = x_2 = x_3$, then $\varepsilon_1 = \varepsilon_2 = 0$, and $\gamma_1 = x^2(\frac{1}{t_2} - \frac{1}{t_1})$, $\gamma_2 = x^2(\frac{1}{t_3} - \frac{1}{t_2})$. The rest parameter estimation procedure is the same as described above.

2-D case: In a homogenous porous medium which is free of solutes initially, the groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$ and $y = \pm\infty$, and will not affect the tracer test results. The analytical solution of BTCs for conservative instantaneous source injected at $x = 0$ and $y = 0$ can be written as Eq. (6). If input mass M is known, there are four parameters (u, n, D_L, D_T) that need to be estimated. We could use four sample concentrations selected at four different times, which are (c_1, t_1, x_1, y_1), (c_2, t_2, x_2, y_2), (c_3, t_3, x_3, y_3), and (c_4, t_4, x_4, y_4) to estimate the four unknown parameters. This is done by substituting above four sets of measurements into Eq. (6).

$$c_1 = \frac{M}{4\pi L n t_1\sqrt{D_L D_T}}\exp\left[-\frac{(x_1-ut_1)^2}{4D_L t_1}\right]\exp\left[-\frac{{y_1}^2}{4D_T t_1}\right] \quad (24)$$

$$c_2 = \frac{M}{4\pi L n t_2\sqrt{D_L D_T}}\exp\left[-\frac{(x_2-ut_2)^2}{4D_L t_2}\right]\exp\left[-\frac{{y_2}^2}{4D_T t_2}\right] \quad (25)$$

$$c_3 = \frac{M}{4\pi L n t_3\sqrt{D_L D_T}}\exp\left[-\frac{(x_3-ut_3)^2}{4D_L t_3}\right]\exp\left[-\frac{{y_3}^2}{4D_T t_3}\right] \quad (26)$$

$$c_4 = \frac{M}{4\pi L n t_4\sqrt{D_L D_T}}\exp\left[-\frac{(x_4-ut_4)^2}{4D_L t_4}\right]\exp\left[-\frac{{y_4}^2}{4D_T t_4}\right] \quad (27)$$

Dividing Eq. (24) by Eq. (25), dividing Eq. (25) by Eq. (26), dividing Eq. (26) by Eq. (27), one can have

$$\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}} = \exp\left[\frac{(x_2-ut_2)^2}{4D_L t_2} - \frac{(x_1-ut_1)^2}{4D_L t_1} + \frac{{y_2}^2}{4D_T t_2} - \frac{{y_1}^2}{4D_T t_1}\right] \quad (28)$$

$$\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}} = \exp\left[\frac{(x_3-ut_3)^2}{4D_L t_3} - \frac{(x_2-ut_2)^2}{4D_L t_2} + \frac{{y_3}^2}{4D_T t_3} - \frac{{y_2}^2}{4D_T t_2}\right] \tag{29}$$

$$\frac{c_3}{c_4}\sqrt{\frac{t_3}{t_4}} = \exp\left[\frac{(x_4-ut_4)^2}{4D_L t_4} - \frac{(x_3-ut_3)^2}{4D_L t_3} + \frac{{y_4}^2}{4D_T t_4} - \frac{{y_3}^2}{4D_T t_3}\right] \tag{30}$$

Eq. (28), Eq. (29), and Eq. (30) could be transformed into Eq. (31), Eq. (32), and Eq. (33) as follows respectively,

$$\frac{1}{4D_T}\left(\frac{{y_1}^2}{t_1} - \frac{{y_2}^2}{t_2}\right) = \frac{(x_2-ut_2)^2}{4D_L t_2} - \frac{(x_1-ut_1)^2}{4D_L t_1} - ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right) \tag{31}$$

$$\frac{1}{4D_T}\left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right) = \frac{(x_3-ut_3)^2}{4D_L t_3} - \frac{(x_2-ut_2)^2}{4D_L t_2} - ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right) \tag{32}$$

$$\frac{1}{4D_T}\left(\frac{{y_3}^2}{t_3} - \frac{{y_4}^2}{t_4}\right) = \frac{(x_4-ut_4)^2}{4D_L t_4} - \frac{(x_3-ut_3)^2}{4D_L t_3} - ln\left(\frac{c_3}{c_4}\sqrt{\frac{t_3}{t_4}}\right) \tag{33}$$

Dividing Eq. (31) by Eq. (32), Dividing Eq. (32) by Eq. (33), one generates Eq. (34) and Eq. (35) as follow,

$$\frac{\left(\frac{{y_1}^2}{t_1} - \frac{{y_2}^2}{t_2}\right)}{\left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right)} = \frac{\frac{1}{4D_L}\left[\frac{(x_2-ut_2)^2}{t_2} - \frac{(x_1-ut_1)^2}{t_1}\right] - ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right)}{\frac{1}{4D_L}\left[\frac{(x_3-ut_3)^2}{t_3} - \frac{(x_2-ut_2)^2}{t_2}\right] - ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right)} \tag{34}$$

$$\frac{\left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right)}{\left(\frac{{y_3}^2}{t_3} - \frac{{y_4}^2}{t_4}\right)} = \frac{\frac{1}{4D_L}\left[\frac{(x_3-ut_3)^2}{t_3} - \frac{(x_2-ut_2)^2}{t_2}\right] - ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right)}{\frac{1}{4D_L}\left[\frac{(x_4-ut_4)^2}{t_4} - \frac{(x_3-ut_3)^2}{t_3}\right] - ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_3}{t_4}}\right)} \tag{35}$$

Eq. (34) and Eq. (35) could be transformed into Eq. (36) and Eq. (37) as follow,

$$\frac{1}{4D_L}(\omega_1 u^2 + \varepsilon_1 u + \gamma_1) = \beta_1 \tag{36}$$

$$\frac{1}{4D_L}(\omega_2 u^2 + \varepsilon_2 u + \gamma_2) = \beta_2 \tag{37}$$

where

$$\omega_1 = \left(\frac{{y_2}^2}{t_2} - \frac{{y_1}^2}{t_1}\right)(t_2 - t_3) + \left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right)(t_1 - t_2),$$

$$\omega_2 = \left(\frac{{y_3}^2}{t_3} - \frac{{y_2}^2}{t_2}\right)(t_3 - t_4) + \left(\frac{{y_3}^2}{t_3} - \frac{{y_4}^2}{t_4}\right)(t_2 - t_3),$$

$$\varepsilon_1 = 2\left[\left(\frac{{y_2}^2}{t_2} - \frac{{y_1}^2}{t_1}\right)(x_3 - x_2) + \left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right)(x_2 - x_1)\right],$$

$$\varepsilon_2 = 2\left[\left(\frac{{y_3}^2}{t_3} - \frac{{y_2}^2}{t_2}\right)(x_4 - x_3) + \left(\frac{{y_3}^2}{t_3} - \frac{{y_4}^2}{t_4}\right)(x_3 - x_2)\right],$$

$$\gamma_1 = \left(\frac{{y_2}^2}{t_2} - \frac{{y_1}^2}{t_1}\right)\left(\frac{x_2^2}{t_2} - \frac{x_3^2}{t_3}\right) + \left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{t_3}\right)\left(\frac{x_1^2}{t_1} - \frac{x_2^2}{t_2}\right),$$

$$\gamma_2 = \left(\frac{{y_3}^2}{t_3} - \frac{{y_2}^2}{t_2}\right)\left(\frac{x_3^2}{t_3} - \frac{x_4^2}{t_4}\right) + \left(\frac{{y_3}^2}{t_3} - \frac{{y_4}^2}{t_4}\right)\left(\frac{x_2^2}{t_2} - \frac{x_3^2}{t_3}\right),$$

$$\beta_1 = \ln\left(\frac{c_1}{c_2}\frac{t_1}{t_2}\right)\left(\frac{{y_3}^2}{t_3} - \frac{{y_2}^2}{t_2}\right) + \ln\left(\frac{c_2}{c_3}\frac{t_2}{t_3}\right)\left(\frac{{y_1}^2}{t_1} - \frac{{y_2}^2}{t_2}\right),$$

$$\beta_2 = \ln\left(\frac{c_2}{c_3}\frac{t_2}{t_3}\right)\left(\frac{{y_4}^2}{t_4} - \frac{{y_3}^2}{t_3}\right) + \ln\left(\frac{c_3}{c_4}\frac{t_3}{t_4}\right)\left(\frac{{y_2}^2}{t_2} - \frac{{y_3}^2}{3}\right).$$

We could divide Eq. (36) by Eq. (37) to get Eq. (38)

$$\frac{\omega_1 u^2 + \varepsilon_1 u + \gamma_1}{\omega_2 u^2 + \varepsilon_2 u + \gamma_2} = \frac{\beta_1}{\beta_2} \tag{38}$$

We could solve Eq. (38) and calculate u

$$u = \frac{(\varepsilon_2\beta_1 - \varepsilon_1\beta_2) + \sqrt{(\varepsilon_1\beta_2 - \varepsilon_2\beta_1)^2 - 4(\omega_1\beta_2 - \omega_2\beta_1)(\gamma_1\beta_2 - \gamma_2\beta_1)}}{2(\omega_1\beta_2 - \omega_2\beta_1)} \tag{39}$$

Then we could substitute u into Eq. (36) for D_L

$$D_L = \frac{\omega_1 u^2 + \varepsilon_1 u + \gamma_1}{4\beta_1} \tag{40}$$

Later substituting u and D_L into Eq. (31) one could have

$$D_T = \frac{\frac{{y_2}^2}{t_2} - \frac{{y_1}^2}{t_1}}{4\ln\left(\frac{c_1 t_1}{c_2 t_2}\right) + \frac{1}{D_L}\left(\frac{(x_1 - ut_1)^2}{4D_L t_1} - \frac{(x_2 - ut_2)^2}{4D_L t_2}\right)} \tag{41}$$

At last substituting D_L, and D_T into Eq. (24) one could have

$$n = \frac{M}{4\pi L c_1 t_1 \sqrt{D_L D_T}} \exp\left[-\frac{(x_1 - ut_1)^2}{4D_L t_1}\right] \exp\left[-\frac{{y_1}^2}{4D_T t_1}\right] \tag{42}$$

This procedure is also applicable if spilled contaminant source mass *M* is unknown. For this case, we has to measure the effective porosity *n* first, then the source mass *M* could be estimated as Eq. (43)

$$M = 4\pi L c_1 t_1 n\sqrt{D_L D_T}\exp\left[\frac{(x_1-ut_1)^2}{4D_L t_1}\right]\exp\left[\frac{{y_1}^2}{4D_T t_1}\right] \quad (43)$$

3-D case: Groundwater flow is along the x-axis with a uniform velocity, and the boundaries are at $x = \pm\infty$, $y = \pm\infty$, and $z = \pm\infty$, and will not affect the tracer test results. The analytical solution of BCTs for conservative instantaneous source injected at $x = 0$, $y = 0$ and $z = 0$ can be written Eq. (7). If input mass M is known, there are five parameters $(u, n, D_L, D_{HT,}\ D_{VT})$ that need to be estimated. We could use five sample concentrations selected at five different times to achieve the objective. Substituting the five data sets of $(c_1, t_1, x_{1,}\ y_1, z_1)$, $(c_2, t_2, x_2, y_2, z_2)$, $(c_3, t_3, x_3, y_3, z_3)$, $(c_4, t_4, x_4, y_4, z_4)$, and $(c_5, t_5, x_5, y_5, z_5)$ into Eq. (7).

$$c_1 = \frac{M}{(4\pi t_1)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}}\exp\left[-\frac{(x_1-ut_1)^2}{4D_L t_1}\right]\exp(-\frac{{y_1}^2}{4D_{HT}t_1})\exp(-\frac{{z_1}^2}{4D_{VT}t_1}) \quad (44)$$

$$c_2 = \frac{M}{(4\pi t_2)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}}\exp\left[-\frac{(x_2-ut_2)^2}{4D_L t_2}\right]\exp(-\frac{{y_2}^2}{4D_{HT}t_2})\exp(-\frac{{z_2}^2}{4D_{VT}t_2}) \quad (45)$$

$$c_3 = \frac{M}{(4\pi t_3)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}}\exp\left[-\frac{(x_3-ut_3)^2}{4D_L t_3}\right]\exp(-\frac{{y_3}^2}{4D_{HT}t_3})\exp(-\frac{{z_3}^2}{4D_{VT}t_3}) \quad (46)$$

$$c_4 = \frac{M}{(4\pi t_4)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}}\exp\left[-\frac{(x_4-ut_4)^2}{4D_L t_4}\right]\exp(-\frac{{y_4}^2}{4D_{HT}t_4})\exp(-\frac{{z_4}^2}{4D_{VT}t_4}) \quad (47)$$

$$c_5 = \frac{M}{(4\pi t_5)^{\frac{3}{2}} n\sqrt{D_L D_{HT} D_{VT}}}\exp\left[-\frac{(x_5-ut_5)^2}{4D_L t_5}\right]\exp(-\frac{{y_5}^2}{4D_{HT}t_5})\exp(-\frac{{z_5}^2}{4D_{VT}t_5}) \quad (48)$$

Dividing Eq. (44) by Eq. (45), dividing Eq. (45) by Eq. (46), dividing Eq. (46) by Eq. (47), and dividing Eq. (47) by Eq. (48), one can have

$$\frac{1}{D_{VT}}\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right) = 4ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}}\right] + \frac{1}{D_L}\left[\frac{(x_1-ut_1)^2}{t_1}\frac{(x_2-ut_2)^2}{t_2}\right] + \frac{1}{D_{HT}}\left(\frac{{y_1}^2}{t_1}-\frac{{y_2}^2}{t_2}\right) \quad (49)$$

$$\frac{1}{D_{VT}}\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right) = 4ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}}\right] + \frac{1}{D_L}\left[\frac{(x_2-ut_2)^2}{t_2}\frac{(x_3-ut_3)^2}{t_3}\right] + \frac{1}{D_{HT}}\left(\frac{{y_2}^2}{t_2}-\frac{{y_3}^2}{t_3}\right) \quad (50)$$

$$\frac{1}{D_{VT}}\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)=4ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}}\right]+\frac{1}{D_L}\left[\frac{(x_3-ut_3)^2}{t_3}\frac{(x_4-ut_4)^2}{t_4}\right]+\frac{1}{D_{HT}}\left(\frac{{y_3}^2}{t_3}-\frac{{y_4}^2}{t_4}\right) \quad (51)$$

$$\frac{1}{D_{VT}}\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)=4ln\left[\frac{c_4}{c_5}\left(\frac{t_4}{t_5}\right)^{\frac{3}{2}}\right]+\frac{1}{D_L}\left[\frac{(x_4-ut_4)^2}{t_4}\frac{(x_5-ut_5)^2}{t_5}\right]+\frac{1}{D_{HT}}\left(\frac{{y_4}^2}{t_4}-\frac{{y_5}^2}{t_5}\right) \quad (52)$$

Dividing Eq. (49) by Eq. (50), dividing Eq. (50) by Eq. (51), dividing Eq. (51) by Eq. (52), one can have

$$\frac{1}{D_{HT}}\left[\left(\frac{{y_2}^2}{t_2}-\frac{{y_3}^2}{t_3}\right)\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)-\left(\frac{{y_1}^2}{t_1}-\frac{{y_2}^2}{t_2}\right)\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\right]=$$

$$4\left\{ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}}\right]\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)-ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}}\right]\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)\right\}+\frac{1}{D_L}\left\{\left(\frac{{z_3}^2}{t_3}-\right.\right.$$

$$\left.\left.\frac{{z_2}^2}{t_2}\right)\left[\frac{(x_1-ut_1)^2}{t_1}\frac{(x_2-ut_2)^2}{t_2}\right]-\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)\left[\frac{(x_2-ut_2)^2}{t_2}\frac{(x_3-ut_3)^2}{t_3}\right]\right\} \quad (53)$$

$$\frac{1}{D_{HT}}\left[\left(\frac{{y_3}^2}{t_3}-\frac{{y_4}^2}{t_4}\right)\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)-\left(\frac{{y_2}^2}{t_2}-\frac{{y_3}^2}{t_3}\right)\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\right]=$$

$$4\left\{ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}}\right]\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)-ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}}\right]\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\right\}+\frac{1}{D_L}\left\{\left(\frac{{z_4}^2}{t_4}-\right.\right.$$

$$\left.\left.\frac{{z_3}^2}{t_3}\right)\left[\frac{(x_2-ut_2)^2}{t_2}\frac{(x_3-ut_3)^2}{t_3}\right]-\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\left[\frac{(x_3-ut_3)^2}{t_3}\frac{(x_4-ut_4)^2}{t_4}\right]\right\} \quad (54)$$

$$\frac{1}{D_{HT}}\left[\left(\frac{{y_4}^2}{t_4}-\frac{{y_5}^2}{t_5}\right)\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)-\left(\frac{{y_3}^2}{t_3}-\frac{{y_4}^2}{t_4}\right)\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)\right]=$$

$$4\left\{ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}}\right]\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)-ln\left[\frac{c_4}{c_5}\left(\frac{t_4}{t_5}\right)^{\frac{3}{2}}\right]\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\right\}+\frac{1}{D_L}\left\{\left(\frac{{z_5}^2}{t_5}-\right.\right.$$

$$\left.\left.\frac{{z_4}^2}{t_4}\right)\left[\frac{(x_3-ut_3)^2}{t_3}\frac{(x_4-ut_4)^2}{t_4}\right]-\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\left[\frac{(x_4-ut_4)^2}{t_4}\frac{(x_5-ut_5)^2}{t_5}\right]\right\} \quad (55)$$

Dividing Eq. (53) by Eq. (54), dividing Eq. (54) by Eq. (55), one can have

$$\frac{\Delta_1}{\Delta_2}=\frac{\varphi_1+\frac{1}{D_L}\left(\Omega_1u^2+b_1u+\emptyset_1\right)}{\varphi_2+\frac{1}{D_L}\left(\Omega_2u^2+b_2u+\emptyset_2\right)} \quad (56)$$

$$\frac{\Delta_2}{\Delta_3}=\frac{\varphi_2+\frac{1}{D_L}\left(\Omega_2u^2+b_2u+\emptyset_2\right)}{\varphi_3+\frac{1}{D_L}\left(\Omega_3u^2+b_3u+\emptyset_3\right)} \quad (57)$$

where $\Delta_1=\left(\frac{{y_2}^2}{t_2}-\frac{{y_3}^2}{t_3}\right)\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)-\left(\frac{{y_1}^2}{t_1}-\frac{{y_2}^2}{t_2}\right)\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)$,

$\Delta_2=\left(\frac{{y_3}^2}{t_3}-\frac{{y_4}^2}{t_4}\right)\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)-\left(\frac{{y_2}^2}{t_2}-\frac{{y_3}^2}{t_3}\right)\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)$,

$\Delta_3=\left(\frac{{y_4}^2}{t_4}-\frac{{y_5}^2}{t_5}\right)\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)-\left(\frac{{y_3}^2}{t_3}-\frac{{y_4}^2}{t_4}\right)\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)$,

$\varphi_1=4\left\{ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}}\right]\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)-ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}}\right]\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)\right\}$,

$\varphi_2=4\left\{ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}}\right]\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)-ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}}\right]\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\right\}$,

$\varphi_3=4\left\{ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}}\right]\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)-ln\left[\frac{c_4}{c_5}\left(\frac{t_4}{t_5}\right)^{\frac{3}{2}}\right]\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\right\}$,

$\Omega_1=\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)t_1+\left(\frac{{z_1}^2}{t_1}-\frac{{z_3}^2}{t_3}\right)t_2+\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)t_3$,

$\Omega_2=\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)t_2+\left(\frac{{z_2}^2}{t_2}-\frac{{z_4}^2}{t_4}\right)t_3+\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)t_4$,

$\Omega_3=\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)t_3+\left(\frac{{z_3}^2}{t_3}-\frac{{z_5}^2}{t_5}\right)t_4+\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)t_5$,

$b_1=2\left[\left(\frac{{z_2}^2}{t_2}-\frac{{z_3}^2}{t_3}\right)x_1+\left(\frac{{z_3}^2}{t_3}-\frac{{z_1}^2}{t_1}\right)x_2+\left(\frac{{z_1}^2}{t_1}-\frac{{z_2}^2}{t_2}\right)x_3\right]$,

$b_2=2\left[\left(\frac{{z_3}^2}{t_3}-\frac{{z_4}^2}{t_4}\right)x_2+\left(\frac{{z_4}^2}{t_4}-\frac{{z_2}^2}{t_2}\right)x_3+\left(\frac{{z_2}^2}{t_2}-\frac{{z_3}^2}{t_3}\right)x_4\right]$,

$b_3=2\left[\left(\frac{{z_4}^2}{t_4}-\frac{{z_5}^2}{t_5}\right)x_3+\left(\frac{{z_5}^2}{t_5}-\frac{{z_3}^2}{t_3}\right)x_4+\left(\frac{{z_3}^2}{t_3}-\frac{{z_4}^2}{t_4}\right)x_5\right]$,

$\emptyset_1=\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\frac{{x_1}^2}{t_1}+\left(\frac{{z_1}^2}{t_1}-\frac{{z_3}^2}{t_3}\right)\frac{{x_2}^2}{t_2}+\left(\frac{{z_2}^2}{t_2}-\frac{{z_1}^2}{t_1}\right)\frac{{x_3}^2}{t_3}$,

$\emptyset_2=\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\frac{{x_2}^2}{t_2}+\left(\frac{{z_2}^2}{t_2}-\frac{{z_4}^2}{t_4}\right)\frac{{x_3}^2}{t_3}+\left(\frac{{z_3}^2}{t_3}-\frac{{z_2}^2}{t_2}\right)\frac{{x_4}^2}{t_4}$,

$\emptyset_3=\left(\frac{{z_5}^2}{t_5}-\frac{{z_4}^2}{t_4}\right)\frac{{x_3}^2}{t_3}+\left(\frac{{z_3}^2}{t_3}-\frac{{z_5}^2}{t_5}\right)\frac{{x_4}^2}{t_4}+\left(\frac{{z_4}^2}{t_4}-\frac{{z_3}^2}{t_3}\right)\frac{{x_5}^2}{t_5}$.

Eq. (56) and Eq. (57) could be transformed to Eq. (58) and Eq. (59), respectively.

$$\frac{1}{D_L}(\omega_1 u^2 + \varepsilon_1 u + \gamma_1) = \beta_1 \tag{58}$$

$$\frac{1}{D_L}(\omega_2 u^2 + \varepsilon_2 u + \gamma_2) = \beta_2 \tag{59}$$

where $\omega_1 = \Delta_1\Omega_2 - \Delta_2\Omega_1$,

$\omega_2 = \Delta_2\Omega_3 - \Delta_3\Omega_2$,

$\varepsilon_1 = \Delta_1 b_2 - \Delta_2 b_1$,

$\varepsilon_2 = \Delta_2 b_3 - \Delta_3 b_2$,

$\gamma_1 = \Delta_1\emptyset_2 - \Delta_2\emptyset_1$,

$\gamma_2 = \Delta_2\emptyset_3 - \Delta_3\emptyset_2$,

$\beta_1 = \Delta_2\varphi_1 - \Delta_1\varphi_2$,

$\beta_2 = \Delta_3\varphi_2 - \Delta_2\varphi_3$.

We could solve Eq. (58) and Eq. (59) to calculate u

$$u = \frac{(\varepsilon_2\beta_1 - \varepsilon_1\beta_2) + \sqrt{(\varepsilon_1\beta_2 - \varepsilon_2\beta_1)^2 - 4(\omega_1\beta_2 - \omega_2\beta_1)(\gamma_1\beta_2 - \gamma_2\beta_1)}}{2(\omega_1\beta_2 - \omega_2\beta_1)} \tag{60}$$

Then we could substitute u into Eq. (58) for D_L

$$D_L = \frac{\omega_1 u^2 + \varepsilon_1 u + \gamma_1}{\beta_1} \tag{61}$$

Later substituting u and D_L into Eq. (53) one could have

$$D_{HT} = \frac{\Delta_1}{4\varphi_1 + \frac{1}{D_L}(\Omega_1 u^2 + b_1 u + \emptyset_1)} \tag{62}$$

Substituting u, D_L and D_{HT} into Eq. (49) one could have

$$D_{VT} = \frac{\left(\frac{{z_2}^2}{t_2} - \frac{{z_1}^2}{t_1}\right)}{4ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}}\right] + \frac{1}{D_L}\left[\frac{(x_1 - ut_1)^2}{t_1} \frac{(x_2 - ut_2)^2}{t_2}\right] + \frac{1}{D_{HT}}\left(\frac{{y_1}^2}{t_1} - \frac{{y_2}^2}{t_2}\right)} \tag{63}$$

At last substituting u, D_L, D_{HT} , and D_{VL} into Eq. (7) one could have

$$n = \frac{M}{(4\pi t_1)^{\frac{3}{2}} c_1 \sqrt{D_L D_{HT} D_{VT}}} \exp\left[-\frac{(x_1 - ut_1)^2}{4D_L t_1}\right] \exp\left(-\frac{{y_1}^2}{4D_{HT} t_1}\right) \exp\left(-\frac{{z_1}^2}{4D_{VT} t_1}\right) \quad (64)$$

This procedure is also applicable if spilled contaminant source mass M is unknown. Similarly, one has to measure the effective porosity n in *a prior* to do so. The source mass M could be estimated as Eq. (65)

$$M = (4\pi t_1)^{\frac{3}{2}} n c_1 \sqrt{D_L D_{HT} D_{VT}} \exp\left[\frac{(x_1 - ut_1)^2}{4D_L t_1}\right] \exp\left(\frac{{y_1}^2}{4D_{HT} t_1}\right) \exp\left(\frac{{z_1}^2}{4D_{VT} t_1}\right) \quad (65)$$

2.2 Reactive Solutes

For reactive solutes, here we consider a linear sorption isotherm and a first-order decay for the tracer transport in porous media with uniform flow fields. Analytical solutions for the 1-D, 2-D, and 3-D cases can be written as Eq. (8), Eq. (9), and Eq. (10), respectively.

The pore velocity, the dispersivities, and the porosity for the reactive solutes in Eq. (8)–(10) can be estimated by the same approach as the conservative solutes. It should be noted in Eq. (8)–(10) that the retardation factor R cannot be estimated by the method of this study. The reason is that R is a rescaling factor of u, i.e., R is lumped with u as a variable u/R (which is the so-called contaminant advective velocity) in Eq. (8)–(10). Thus, the peak time t_m only contains the information of the lumped variable u/R but not individual u or R. Nevertheless, R can be estimated easily by comparing the peak time of a reactive solute with the peak time of a conservative solute for the same porous media if necessary. λ can be estimated based on the method of this study by adding one more observation point. However, λ is mainly determined by the chemical characteristics of the solute itself such as the half-life. Thus, the λ estimation

is not discussed here. When the retardation factor R and the reaction rate λ are obtained in advance, the procedures of estimating the rest parameters are outlined as follows.

1-D case: now three observed datasets (c_1, t_1, x_1), (c_2, t_2, x_2), and (c_3, t_3, x_3) should be substituted into Eq. (8) to calculate three parameters (u, n, α_L).

$$c_1 = \frac{M}{2AnR\sqrt{D_L\pi t_1/R}} \exp\left[-\frac{(x_1-ut_1/R)^2}{4D_L t_1/R} - \lambda t_1\right] \tag{66}$$

$$c_2 = \frac{M}{2AnR\sqrt{D_L\pi t_2/R}} \exp\left[-\frac{(x_2-ut_2/R)^2}{4D_L t_2/R} - \lambda t_2\right] \tag{67}$$

$$c_3 = \frac{M}{2AnR\sqrt{D_L\pi t_3/R}} \exp\left[-\frac{(x_3-ut_3/R)^2}{4D_L t_3/R} - \lambda t_3\right] \tag{68}$$

Dividing Eq. (66) by Eq. (67), and Eq. (67) by Eq. (68), one can have

$$\ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right) = \frac{(x_2-ut_2/R)^2}{4D_L t_2/R} - \frac{\left(x_1-\frac{ut_1}{R}\right)^2}{\frac{4D_L t_1}{R}} + \lambda(t_2 - t_1) \tag{69}$$

$$\ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right) = \frac{(x_3-ut_3/R)^2}{4D_L t_3/R} - \frac{\left(x_2-\frac{ut_2}{R}\right)^2}{\frac{4D_L t_1}{R}} + \lambda(t_3 - t_2) \tag{70}$$

Dividing Eq. (69) by Eq. (70), we could have the core equation for pore velocity

$$(\omega_1\beta_2 - \omega_2\beta_1)u^2\text{-}(\varepsilon_1\beta_2 - \varepsilon_2\beta_1)u\text{+}(\gamma_1\beta_2 - \gamma_2\beta_1) = 0 \tag{71}$$

where $\omega_1 = \frac{t_2-t_1}{R^2}$, $\varepsilon_1 = \frac{2(x_1-x_2)}{R}$, $\gamma_1 = \frac{x_2^2}{t_2} - \frac{x_1^2}{t_1}$, $\beta_1 = \ln\left(\frac{c_1}{c_2}\sqrt{\frac{t_1}{t_2}}\right) + \lambda(t_1 - t_2)$,

$\omega_2 = \frac{t_3-t_2}{R^2}$, $\varepsilon_2 = \frac{2(x_2-x_3)}{R}$, $\gamma_2 = \frac{x_3^2}{t_3} - \frac{x_2^2}{t_2}$, $\beta_2 = \ln\left(\frac{c_2}{c_3}\sqrt{\frac{t_2}{t_3}}\right) + \lambda(t_2 - t_3)$.

We could solve Eq. (71) and calculate u

$$u = \frac{(\varepsilon_2\beta_1-\varepsilon_1\beta_2)+\sqrt{(\varepsilon_1\beta_2-\varepsilon_2\beta_1)^2-4(\omega_1\beta_2-\omega_2\beta_1)(\gamma_1\beta_2-\gamma_2\beta_1)}}{2(\omega_1\beta_2-\omega_2\beta_1)} \tag{72}$$

Then we could substitute u into Eq. (69) for D_L

$$D_L = \frac{R(\omega_1 u^2+\varepsilon_1 u+\gamma_1)}{4\beta_1} \tag{73}$$

Later substituting u and D_L into Eq. (66) one could have

$$n = \frac{M}{2Anc_1R\sqrt{D_L\pi t_1/R}}\exp\left[-\frac{(x_1-ut_1/R)^2}{4D_Lt_1/R}-\lambda t_1\right] \tag{74}$$

This procedure is also applicable if spilled contaminant source mass *M* is unknown. The source mass *M* could be estimated as Eq. (75) if the effective porosity *n* is determined in advance:

$$M = 2Ac_1\, nR\sqrt{D_L\pi t_1/R}\exp\left[\frac{(x_1-ut_1/R)^2}{4D_Lt_1/R}+\lambda t_1\right] \tag{75}$$

Similarly, the results for 2-D case are as follows:

$$u = \frac{(\varepsilon_2\beta_1-\varepsilon_1\beta_2)+\sqrt{(\varepsilon_1\beta_2-\varepsilon_2\beta_1)^2-4(\omega_1\beta_2-\omega_2\beta_1)(\gamma_1\beta_2-\gamma_2\beta_1)}}{2(\omega_1\beta_2-\omega_2\beta_1)} \tag{76}$$

$$D_L = \frac{R(\omega_1u^2+\varepsilon_1u+\gamma_1)}{4\beta_1} \tag{77}$$

$$D_T = \frac{R\Delta_1}{4\varphi_1+\frac{R}{D_L}(\Omega_1u^2+b_1u+\emptyset_1)} \tag{78}$$

$$n = \frac{M}{4\pi Lc_1t_1\sqrt{D_LD_T}}\exp\left[-\frac{(x_1-ut_1/R)^2}{4D_Lt_{1/R}}-\lambda t_1\right]\exp\left[-\frac{{y_1}^2}{4D_Tt_1/R}\right] \tag{79}$$

$$M = 4\pi Lc_1t_1n\sqrt{D_LD_T}\exp\left[\frac{(x_1-ut_1/R)^2}{4D_Lt_{1/R}}+\lambda t_1\right]\exp\left[\frac{{y_1}^2}{4D_Tt_1}\right] \tag{80}$$

where $\Delta_1=\frac{{y_2}^2}{t_2}-\frac{{y_1}^2}{t_1}$,

$\Delta_2=\frac{{y_3}^2}{t_3}-\frac{{y_2}^2}{t_2}$,

$\Delta_3=\frac{{y_4}^2}{t_4}-\frac{{y_3}^2}{t_3}$,

$\varphi_1 = ln\left(\frac{c_1}{c_2}\frac{t_1}{t_2}\right)+\lambda(t_1-t_2)$,

$\varphi_2 = ln\left(\frac{c_2}{c_3}\frac{t_2}{t_3}\right)+\lambda(t_2-t_3)$,

$\varphi_3 = ln\left(\frac{c_2}{c_3}\frac{t_2}{t_3}\right)+\lambda(t_3-t_4)$,

$$\Omega_1 = \frac{t_1 - t_2}{R^2},$$

$$\Omega_2 = \frac{t_2 - t_3}{R^2},$$

$$\Omega_3 = \frac{t_3 - t_4}{R^2},$$

$$b_1 = \frac{2(x_2 - x_1)}{R},$$

$$b_2 = \frac{2(x_3 - x_2)}{R},$$

$$b_3 = \frac{2(x_4 - x_3)}{R},$$

$$\emptyset_1 = \frac{{x_1}^2}{t_1} - \frac{{x_2}^2}{t_2},$$

$$\emptyset_2 = \frac{{x_2}^2}{t_2} - \frac{{x_3}^2}{t_3},$$

$$\emptyset_3 = \frac{{x_3}^2}{t_3} - \frac{{x_4}^2}{t_4}.$$

$$\omega_1 = \Delta_1\Omega_2 - \Delta_2\Omega_1,$$

$$\omega_2 = \Delta_2\Omega_3 - \Delta_3\Omega_2,$$

$$\varepsilon_1 = \Delta_1 b_2 - \Delta_2 b_1,$$

$$\varepsilon_2 = \Delta_2 b_3 - \Delta_3 b_2,$$

$$\gamma_1 = \Delta_1\emptyset_2 - \Delta_2\emptyset_1,$$

$$\gamma_2 = \Delta_2\emptyset_3 - \Delta_3\emptyset_2,$$

$$\beta_1 = \Delta_2\varphi_1 - \Delta_1\varphi_2,$$

$$\beta_2 = \Delta_3\varphi_2 - \Delta_2\varphi_3.$$

And the results for 3-D case are as follows:

$$u = \frac{(\varepsilon_2\beta_1 - \varepsilon_1\beta_2) + \sqrt{(\varepsilon_1\beta_2 - \varepsilon_2\beta_1)^2 - 4(\omega_1\beta_2 - \omega_2\beta_1)(\gamma_1\beta_2 - \gamma_2\beta_1)}}{2(\omega_1\beta_2 - \omega_2\beta_1)} \tag{81}$$

$$D_L = \frac{R(\omega_1 u^2 + \varepsilon_1 u + \gamma_1)}{4\beta_1} \tag{82}$$

$$D_{HT} = \frac{R\Delta_1}{4\varphi_1 + \frac{R}{D_L}(\Omega_1 u^2 + b_1 u + \emptyset_1)} \tag{83}$$

$$D_{VT} = \frac{R\left(\frac{z_2^2}{t_2} - \frac{z_1^2}{t_1}\right)}{4ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}}\right] + 4\lambda(t_1 - t_2) + \frac{(x_1 - ut_1/R)^2}{4D_L t_1/R} - \frac{(x_2 - ut_2/R)^2}{4D_L t_2/R} + \frac{y_1^2}{4D_{HT} t_1/R} - \frac{y_2^2}{4D_{HT} t_2/R}} \tag{84}$$

$$n = \frac{M}{(4\pi t_1)^{3/2} c_1 \sqrt{D_L D_{HT} D_{VT}}} \exp\left[-\frac{(x_1 - ut_1/R)^2}{4D_L t_1/R} - \lambda t_1\right] \exp\left(-\frac{y_1^2}{4D_{HT} t_1/R}\right) \exp\left(-\frac{z_1^2}{4D_{VT} t_1/R}\right) \tag{85}$$

$$M = (4\pi t_1)^{3/2} c_1 n \sqrt{D_L D_{HT} D_{VT}} \exp\left[\frac{(x_1 - u^* t_1/R)^2}{4D_L t_{1/R}} + \lambda t_1\right] \exp\left(\frac{y_1^2}{4D_{HT} t_{1/R}}\right) \exp\left(\frac{z_1^2}{4D_{VT} t_{1/R}}\right) \tag{86}$$

where $\Delta_1 = \left(\frac{y_2^2}{t_2} - \frac{y_3^2}{t_3}\right)\left(\frac{z_2^2}{t_2} - \frac{z_1^2}{t_1}\right) - \left(\frac{y_1^2}{t_1} - \frac{y_2^2}{t_2}\right)\left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right)$,

$\Delta_2 = \left(\frac{y_3^2}{t_3} - \frac{y_4^2}{t_4}\right)\left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right) - \left(\frac{y_2^2}{t_2} - \frac{y_3^2}{t_3}\right)\left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right)$,

$\Delta_3 = \left(\frac{y_4^2}{t_4} - \frac{y_5^2}{t_5}\right)\left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right) - \left(\frac{y_3^2}{t_3} - \frac{y_4^2}{t_4}\right)\left(\frac{z_5^2}{t_5} - \frac{z_4^2}{t_4}\right)$,

$\varphi_1 = ln\left[\frac{c_1}{c_2}\left(\frac{t_1}{t_2}\right)^{\frac{3}{2}} + \lambda(t_1 - t_2)\right]\left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right) - ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}} + \lambda(t_2 - t_3)\right]\left(\frac{z_2^2}{t_2} - \frac{z_1^2}{t_1}\right)$,

$\varphi_2 = ln\left[\frac{c_2}{c_3}\left(\frac{t_2}{t_3}\right)^{\frac{3}{2}} + \lambda(t_2 - t_3)\right]\left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right) - ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}} + \lambda(t_3 - t_4)\right]\left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right)$,

$\varphi_3 = ln\left[\frac{c_3}{c_4}\left(\frac{t_3}{t_4}\right)^{\frac{3}{2}} + \lambda(t_3 - t_4)\right]\left(\frac{z_5^2}{t_5} - \frac{z_4^2}{t_4}\right) - ln\left[\frac{c_4}{c_5}\left(\frac{t_4}{t_5}\right)^{\frac{3}{2}} + \lambda(t_4 - t_5)\right]\left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right)$,

$\Omega_1 = \frac{1}{R^2}\left[\left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right)(t_1 - t_2) - \left(\frac{z_2^2}{t_2} - \frac{z_1^2}{t_1}\right)(t_2 - t_3)\right]$,

$\Omega_2 = \frac{1}{R^2}\left[\left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right)(t_2 - t_3) - \left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right)(t_2 - t_4)\right]$,

$\Omega_3 = \frac{1}{R^2}\left[\left(\frac{z_5^2}{t_5} - \frac{z_4^2}{t_4}\right)(t_3 - t_4) - \left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right)(t_3 - t_5)\right]$,

$b_1 = \frac{2}{R}\left[\left(\frac{z_3^2}{t_3} - \frac{z_1^2}{t_1}\right)(x_2 - x_1) - \left(\frac{z_2^2}{t_2} - \frac{z_1^2}{t_1}\right)(x_3 - x_2)\right]$,

$b_2 = \frac{2}{R}\left[\left(\frac{z_4^2}{t_4} - \frac{z_2^2}{t_2}\right)(x_3 - x_2) - \left(\frac{z_3^2}{t_3} - \frac{z_2^2}{t_2}\right)(x_4 - x_3)\right]$,

$b_3 = \frac{2}{R}\left[\left(\frac{z_5^2}{t_5} - \frac{z_3^2}{t_3}\right)(x_4 - x_3) - \left(\frac{z_4^2}{t_4} - \frac{z_3^2}{t_3}\right)(x_5 - x_4)\right]$,

$$\emptyset_1 = \left(\frac{{z_3}^2}{t_3} - \frac{{z_2}^2}{t_2}\right)\left(\frac{{x_1}^2}{t_1} - \frac{{x_2}^2}{t_2}\right) - \left(\frac{{z_2}^2}{t_2} - \frac{{z_1}^2}{t_1}\right)\left(\frac{{x_2}^2}{t_2} - \frac{{x_3}^2}{t_3}\right),$$

$$\emptyset_2 = \left(\frac{{z_4}^2}{t_4} - \frac{{z_3}^2}{t_3}\right)\left(\frac{{x_2}^2}{t_2} - \frac{{x_3}^2}{t_3}\right) - \left(\frac{{z_3}^2}{t_3} - \frac{{z_2}^2}{t_2}\right)\left(\frac{{x_3}^2}{t_3} - \frac{{x_4}^2}{t_4}\right),$$

$$\emptyset_3 = \left(\frac{{z_5}^2}{t_5} - \frac{{z_4}^2}{t_4}\right)\left(\frac{{x_3}^2}{t_3} - \frac{{x_4}^2}{t_4}\right) - \left(\frac{{z_4}^2}{t_4} - \frac{{z_3}^2}{t_3}\right)\left(\frac{{x_4}^2}{t_4} - \frac{{x_5}^2}{t_5}\right),$$

$$\omega_1 = \Delta_1\Omega_2 - \Delta_2\Omega_1,$$

$$\omega_2 = \Delta_2\Omega_3 - \Delta_3\Omega_2,$$

$$\varepsilon_1 = \Delta_1 b_2 - \Delta_2 b_1,$$

$$\varepsilon_2 = \Delta_2 b_3 - \Delta_3 b_2,$$

$$\gamma_1 = \Delta_1\emptyset_2 - \Delta_2\emptyset_1,$$

$$\gamma_2 = \Delta_2\emptyset_3 - \Delta_3\emptyset_2,$$

$$\beta_1 = \Delta_2\varphi_1 - \Delta_1\varphi_2,$$

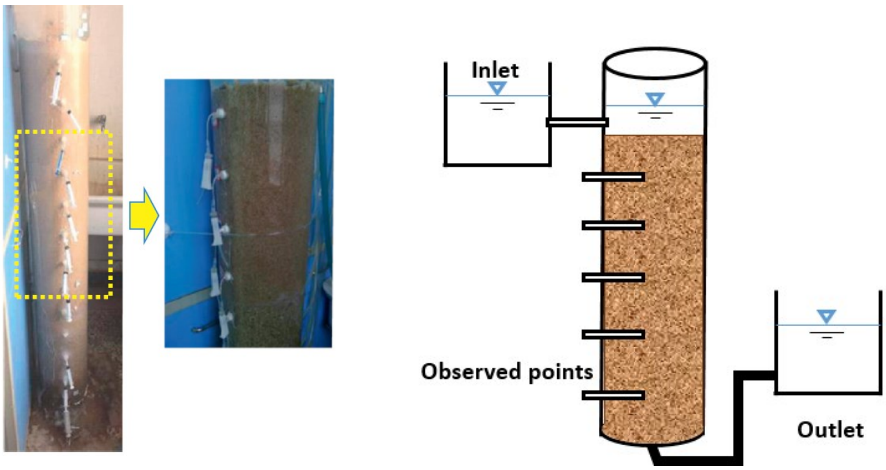
$$\beta_2 = \Delta_3\varphi_2 - \Delta_2\varphi_3.$$

3. LABORATORY EXPERIMENTS

3.1 Settings for four cases

Dr. Guiming Dong carried out a batch of column tests in School of Resources and Geosciences, China University of Mining and Technology in the year of 2017. In this study, I collected some of his experimental data to analyze whether the proposed method could be used to obtain creditable estimation for the parameters in the laboratory experiments. Based on these data I set up two cases for comparison to analyze whether the estimation effect is related to the observation point coordinates and flow velocity or not. A column filled with sand is constructed by a Perspex pipe (Figure 1.). The steady flow is maintained by setting constant head boundaries at the inlet and outlet of the sand columns. Before injecting tracer tests, the porosity of the column medium is determined as the laboratory porosity described below, and the average hydraulic conductivity is estimated by the Darcy experiments. The tests are implemented in two sand columns with two different uniform flow velocities. The two columns contain homogeneous medium to coarse sands whose diameters are around 0.5–0.7mm. The diameter of the two columns is 19cm and the length is 1.3m. The average hydraulic conductivity of coarse sands is 3.72cm/min (53.63m/d), estimated by the Darcy's experiments discussed in details below. The 0.33L chloride with a concentration of 1690mg/L is instantly injected into the inlet column. The chloride concentrations are measured at #4 observed point x_1=30cm and #8 observed point x_2=70cm from the inlet. The time interval of the monitor is around 20min. The tracer tests are implemented with Darcy velocities of 0.10cm/min and 0.05cm/min in these two columns.

The sample volume for measuring tracer concentration is so small that it does not affect the flow velocity in the columns. Chloride is adopted as a conservative tracer. The sodium chloride solution is instantaneously injected into the columns at the inlet. The concentration of chloride is obtained by measuring the conductivity of the water sample using the transducer of Solinst 3001 Levelogger® Edge that has a measured range of 0–80,000 *μs/cm* and an accuracy of 0.05% FS. This transducer is ideal for salinity and saltwater intrusion studies and tracer tests.



A B

Figure 1. (A) The picture of the sand column of the laboratory injecting tracer tests. (B) The schematic diagram for the laboratory injecting tracer tests.

The chloride concentration can be obtained directly from the measured conductivity based on a linear relation between the conductivity and chloride concentration, which has been calibrated right before the experiment. Table 1 below

demonstrates how the calibration experiment is designed. Specific amount of tracer NaCl would be added into 7 liters distilled water 23 times sequentially. After dissolution one can calculate the concentration of NaCl or Cl^-, measure the conductivity and current temperature using the transducer at each time. Later we could use Eq. (87) to obtain the conductivity of solution at reference temperature (25 °C), what is termed as compensation conductivity. Through a linear match between total Cl^- concentration (mg/L) and 25 °C conductivity (μS/cm), one could find the linear correlation for the purpose of best fitting, like what is obtained in Figure 2. Based on the calibration criteria, we could calculate Cl^- concentration at observed point by testing the conductivity at specific time.

The temperature compensation coefficient during this time is set as α = 2.00 %/°C. The measured conductivity at reference temperature (here 25 °C) should be

$$G_{TRef}=\frac{G_T}{\frac{\alpha(T-T_{Ref})}{100\%/°C}+1} \tag{87}$$

where G_{TRef} is measured conductivity at reference temperature, G_T is measured conductivity at temperature T, T_{Ref} –reference temperature (25 °C), and α is temperature compensation coefficient (2.00 %/°C).

Table 1. Calibration experiment designed for linear correlation between Cl^- concentration (mg/L) and compensation conductivity (μS/cm).

ID	NaCl each time (g)	Total NaCl (g)	Total Cl^- (mg/L)	Conductivity (μS/cm)	Temperature (℃)	25℃ Conductivity (μS/cm)
0	0.0000	0.0000	0.000	688	13.1	902.9
1	0.0560	0.0560	4.853	714	12.9	942.0
2	0.0560	0.1120	9.706	725	12.9	956.5
3	0.0600	0.1720	14.906	738	12.9	973.6
4	0.0583	0.2303	19.958	750	12.8	992.1
5	0.0630	0.2933	25.418	762	12.8	1007.9
6	0.0540	0.3473	30.097	773	12.8	1022.5
7	0.0597	0.4070	35.271	785	12.7	1041.1
8	0.0654	0.4724	40.939	797	12.7	1057.0
9	0.0625	0.5349	46.355	810	12.7	1074.3
10	0.5015	1.0364	89.816	908	12.7	1204.2
11	0.9412	1.9776	171.381	1090	12.6	1449.5
12	1.3592	3.3368	289.171	1350	12.6	1795.2
13	1.6809	5.0177	434.840	1699	12.5	2265.3
14	2.7003	7.7180	668.851	2257	12.5	3009.3
15	3.7165	11.4345	990.927	3026	12.5	4034.7
16	4.3693	15.8038	1369.576	3928	12.5	5237.3
17	4.8227	20.6265	1787.517	4922	12.3	6597.9
18	5.4883	26.1148	2263.140	6047	12.2	8127.7
19	6.3972	32.5120	2817.529	7356	12.2	9887.1

Table 1 continued

ID	NaCl each time (g)	Total NaCl (g)	Total Cl- (mg/L)	Conductivity (μS/cm)	Temperature (℃)	25℃ Conductivity (μS/cm)
20	7.4734	39.9854	3465.183	8858	12.1	11938.0
21	8.5492	48.5346	4206.066	10522	12.1	14180.6
22	9.7415	58.2761	5050.277	12386	12.0	16737.8
23	10.1214	68.3975	5927.409	14345	12.0	19385.1

* Total volume for calibration is 7L

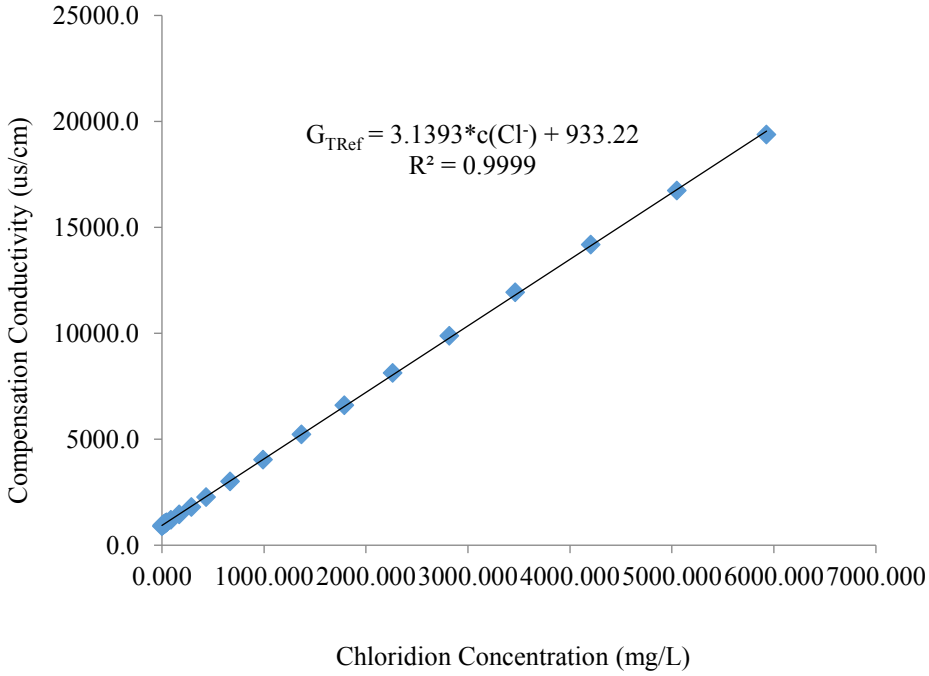


Figure 2. The linear correlation between conductivity at reference temperature (25 °C) and Cl^- concentration.

The porosity of material is the percentage of the medium that is void of material. Laboratory porosity is determined by taking a sample of known volume. The sample is dried in an oven at 105 ℃ until it reaches a constant weight. This expels moisture clinging to surfaces in the sample, but not water that is hydrated as a part of certain minerals. The dried sample is then submerged in a known volume of water and allowed to remain in a sealed chamber until it is saturated. The volume of the voids is equal to the original water volume less the volume in the chamber after the saturated sample is removed. Dr. Guiming Dong concluded that the porosity of the coarse sands is 0.32.

Hydraulic conductivity (K) is a measure of a material's capacity to transmit water, which is a function of water viscosity and density, both are functions of water temperature. However, given the small range of temperature variation encountered in most groundwater systems, the temperature dependence of hydraulic conductivity is often neglected. One can employ Darcy's Law to estimate the hydraulic conductivity K $[L/T]$ of two types of porous media in the columns. The equation for Darcy's Law is based on the observations that the flow rate through a porous medium (such as an aquifer) is proportional to the cross-sectional area perpendicular to flow and is also proportional to the head loss per unit length in the direction of flow. Putting these two proportionalities together gives the following equation:

$$Q = -KA\frac{dh}{dl} \tag{88}$$

where Q is flow rate of solution through the porous medium [L^3/T], one can use measuring cylinder and stopwatch to get the volume of outflow in specific time period to obtain the flow rate, K is hydraulic conductivity of the porous media [L/T], A is

cross-sectional area perpendicular to flow [L^2], *dh/dl* is hydraulic gradient [dimensionless], the head would decrease in the direction of flow, so negative sign means the flow direction is from the high head to the low head. Table 2 illustrates how to design parallel tests for *K* value estimation in coarse sands (0.5–0.7 mm).

The diameter of the column is 19cm and the total length is 1.3m, so the cross-section area is 283.39cm^2. Table 3 lists the result of *K* value, and the average hydraulic conductivity of coarse sands is 3.72cm/min (53.63m/d).

Table 2. Test under the flow rate of 7.068ml/sec in coarse sand column.

Time (sec)	Volume (ml)	Flow (ml/sec)	Observed Point # *x*-distance from inlet	Head (cm)
44	311	7.068	1 (*x*=5cm)	76.3
			2 (*x*=10cm)	74.6
			3 (*x*=20cm)	70.1
			4 (*x*=30cm)	65.9
			6 (*x*=50cm)	57.6

Table 3**.** *K* value under the flow rate of 7.068ml/sec in coarse sands column.

A (cm^2)	ΔH_1 (cm)	*L* (cm)	J_1	K_1 (cm/s)	K_1 (m/d)
283.385	1.7	5	0.34	0.073	63.382
	4.5	10	0.45	0.055	47.889
	4.2	10	0.42	0.059	51.309
	8.3	20	0.415	0.060	51.927
Average *K* value (m/d)				53.627	

3.2 Output of injecting tracer tests

Case 1: The column is filled with homogeneous coarse sand whose diameters are around 0.5–0.7mm (coarse sand), the velocity is 0.10cm/min, and the linear correlation between conductivity and chloride concentration is calibrated both as $G_{TRef} = 2.9549*c(Cl^-) + 950.83$ for observed points at x_1=30cm and x_2=70cm. Laboratory data are attached in Appendix A-1.

Case 2: The column is filled with homogeneous coarse sand whose diameters are around 0.5–0.7mm (coarse sand), the velocity is 0.05 cm/min, and the linear correlation between conductivity and chloride concentration is calibrated both as $G_{TRef} = 2.9549*c(Cl^-) + 950.83$ for observed points at x_1=30cm and x_2=70cm. Laboratory data are attached in Appendix A-2.

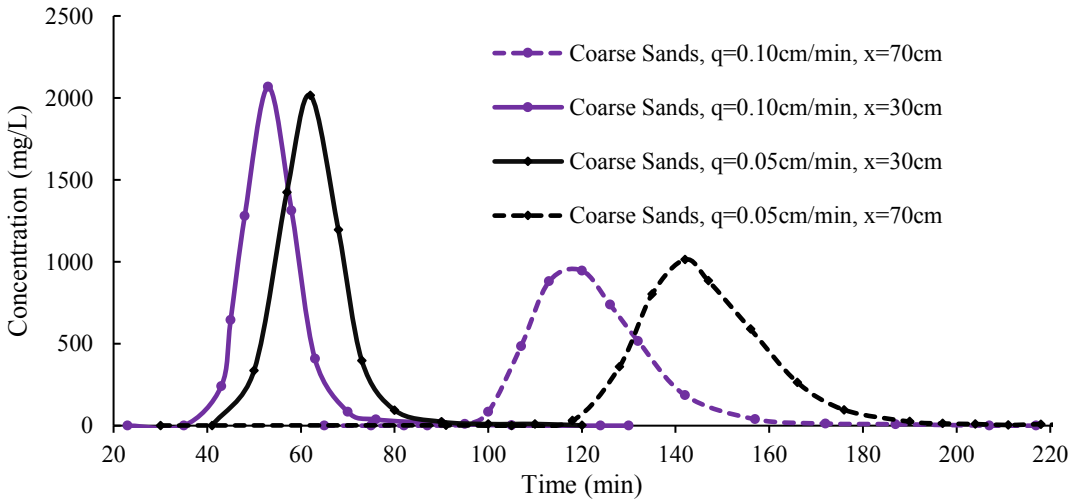


Figure 3. The breakthrough curves at different observed points in two columns with two different velocity.

Figure 3 shows that the tracer concentration tested at different observed points in different porous media with different velocity, their BTCs all follow somewhat a bell shape, but certainly not the normal distribution as one can clearly see the skewness of those curves: they all have a steeper climbing limb but a gentle declining limb.

The Liang's method concludes that the peak time, t_m, could be written as follow

$$t_m = -\frac{\alpha_L}{u} + \sqrt{\left(\frac{\alpha_L}{u}\right)^2 + \left(\frac{x}{u}\right)^2} \tag{89}$$

Eq. (89) shows that the peak time depends on the location of the observation point along the longitudinal direction. If the observation points are farther away from the inlets, it will take them longer time for the peak concentration to appear, and the peak concentration would be lower. As it is shown that solid curves have the higher (twice larger) peak concentration tested at earlier peak time (t_m) than the dashed curves in the same color.

The peak time is also controlled by the flow rate: the faster the velocity is, the earlier the peak times is observed. However, the peak concentration is independent of the flow rate. Here in the coarse sand columns, peak time is obtained 10 minutes earlier at observation point of x=30cm with a velocity of 0.10cm/min (peak time is 52 minutes in blue solid curve) than the one obtained with a velocity of 0.05cm/min (peak time is 62minutes in black solid curve). Such delay of peak time is more obvious at the observation points near outlets (x=70cm), where the peak time is 120minutes with a velocity of 0.10cm/min and 150minutes with a lower velocity of 0.05cm/min. Such a 30minutes delay could be found between blue dashed curve and the black dashed curve. However the peak concentrations at certain observation points have not varied considerably during this delay period.

4. RESULTS AND DISCUSSION

Use 1-D conservative instantaneous source test for the generalized model, we should choose three sets of sample data (at the same observed point) to estimate three unknown parameters (u, n, α_L), here three sets of sample data (t_1, C_1), (t_2, C_2), and (t_3, C_3) are obtained at three different transport time (t_1, t_2, t_3), and corresponding observed concentrations (C_1, C_2, C_3). To serve this purpose, I have programed MATLAB script files for parameter estimation for all the possible combination of data sets. Specifically, if there are *N* sets of data $((t_1, C_1), (t_2, C_2),\ldots,(t_N, C_N))$ at a certain observed point in the laboratory column test, we should consider cases of total C_N^3 combination, explained as follow

$$C_N^3 = \frac{N!}{3!(N-3)!} \tag{90}$$

where $k!$ is k factorial and it is defined by $N! = N \times (N-1) \times \cdots \times 3 \times 2 \times 1$, $(N-3)!$ is $(N-3)$ factorial defined by $(N-3)! = (N-3) \times (N-4) \times \cdots \times 3 \times 2 \times 1$, and $3! = 3 \times 2 \times 1 = 6$.

There is one thing we should notice, not all these C_N^3 combinations could be used to calculate the parameters, because some pairs of data would make the coefficients not creditable, so the MATLAB script programmed in Appendix B-1 would only count the number of combinations which could make the coefficients calculable. For each combination, we could obtain parameters (u, n, α_L), then calculate the root-mean-square-error (RMSE) after substituting the parameters into the generalized model and comparing with real laboratory data. All this work should be done for the purpose of selecting the optimized sampling time.

In addition, I have programed the MATLAB script files for CFM to estimate parameters based on the principle of minimum RMSE, and for parameter estimation using the Liang's method as well. At last, I will compare these two methods with the proposed new method of this thesis, and to conduct an uncertainty analysis.

In the following analysis, we will analyze the two cases discussed in the experimental studies of Chapter 3.

4.1 Estimation and comparison for case 1

Case 1: q=0.10 cm/min, sand diameters are around 0.5–0.7mm, K=3.72cm/min or 53.63m/d.

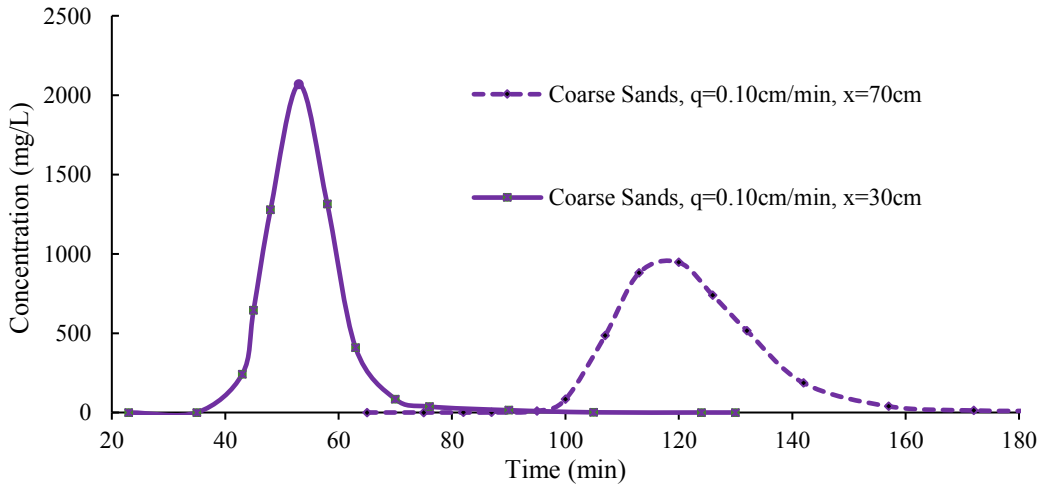


Figure 4. The breakthrough curves at two observation points in coarse sand column with a velocity of 0.10cm/min.

Firstly we focus on the observed point at x_1=30cm, as listed in Appendix A-1, we collected 13 pairs ($N = 13$) of data ((t_1, C_1), (t_2, C_2),…,(t_{13}, C_{13})), one can pick three pairs of observed times and corresponding concentrations ((t_i, C_i), (t_j, C_j),

(t_k, C_k), $1 \le i < j < k \le 13$). Then one can use MATLAB script file in Appendix B-1 to obtain the final 91 combination sets of three different observed data pairs.

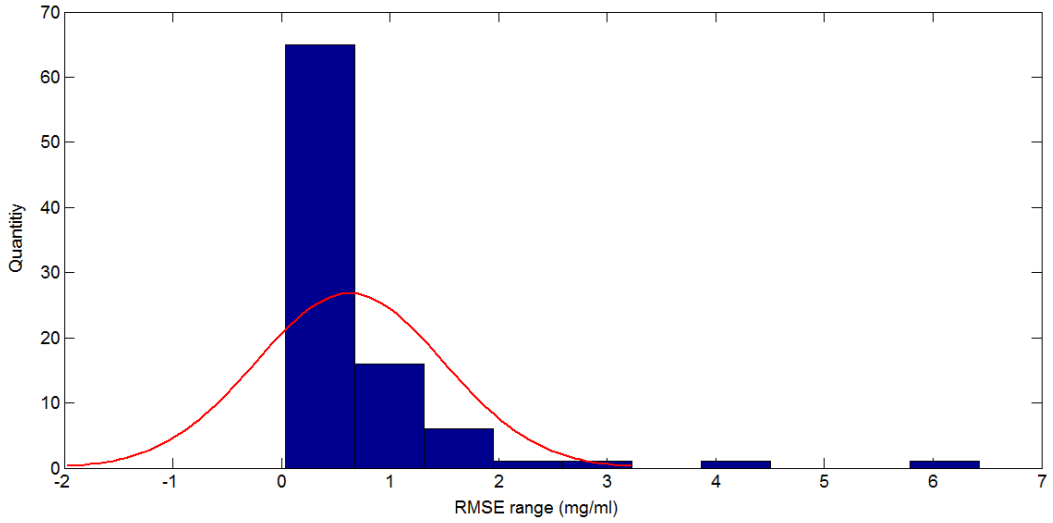


Figure 5. Histogram for all possibilities at *x*=30cm in case 1.

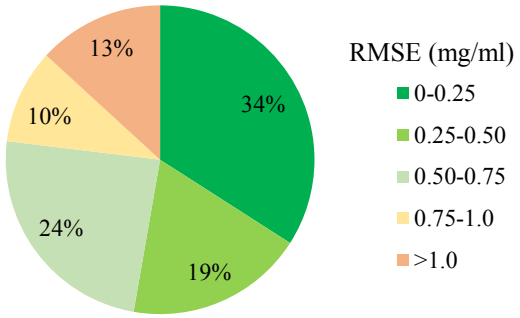


Figure 6. RMSE relative frequency distribution chart at *x*=30cm in case 1.

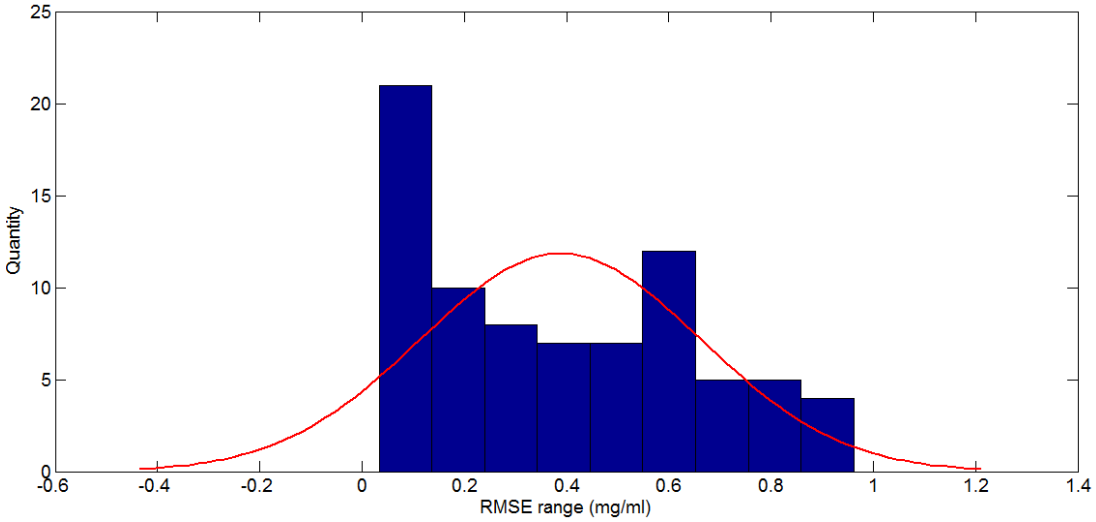


Figure 7. Histogram for only RMSE<1mg/ml at x=30cm in case 1.

The RMSE ranges from 0.0334mg/ml (1.62% of c_m=2066mg/L) to 6.4241mg/ml, the minimum RMSE occurs when we choose t_1=43min (c_1=241mg/L), t_2=53min (c_2=c_m=2066mg/L), and t_3=58min (c_3=1313mg/L), while in this case, the peak concentration c_m of 2066mg/L is tested at t_m=53min. So for this situation, the best selection for the smallest RMSE requires that one chooses the time right after the solute passes through this observed point of x_1=30cm. On the contrary, if one chooses t_1=43min (c_1=241mg/L), t_2=45min (c_2=643mg/L), and t_3=105min (c_3=1mg/L), the deviation from actual data would be the greatest with the largest RMSE. This happens when we select all three time points at the very end of the curve decreasing limb.

The RMSE frequency distribution could be analyzed by Figure 5, 6, and 7. As can be seen from these figures, 34% of the total 91 combination selection has 0-0.25mg/ml RMSE (which corresponds to 0%-12.10% of c_m) using the parameters estimated by the proposed method. Similarly, 19% of sets have the result of RMSE above 0.25mg/ml but below 0.50mg/ml (which corresponds to 12.10%-24.20% of c_m),

24% of sets creates the RMSE higher than 0.50mg/ml lower than 0.75mg/ml (which corresponds to 24.20%- 36.30% of c_m), the rest 23% of sets make the result deviated more from the real data with the RMSE greater than 0.75mg/ml.

Then if we focus on the time sets corresponding to the smallest RMSE of 0.0334mg/ml (1.62% of c_m), the parameters could be estimated as: n=0.13, u=0.565 cm/min, and D_L=0.083cm^2/min. If the effective molecular coefficient is neglected, then α_L=0.148cm. Figure6 below is created using the MATLAB script in Appendix B-2, and it shows that BTC at the observed point x_1=30cm with the parameters estimated by the proposed method fits well with the laboratory data.

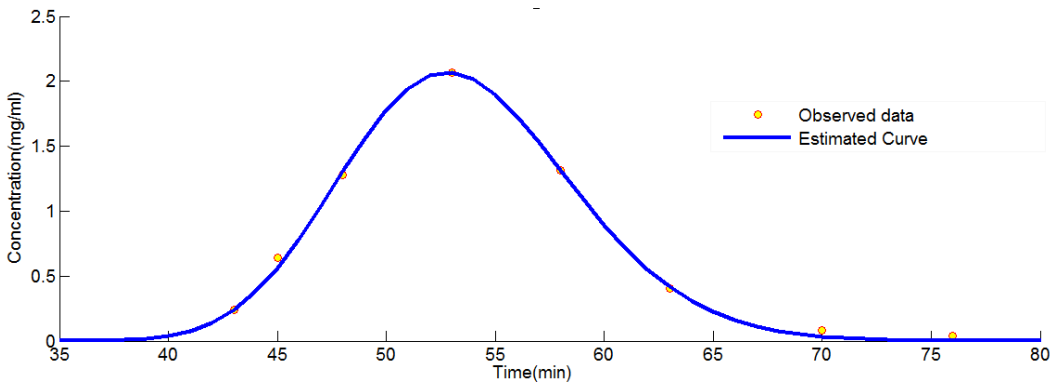


Figure 8. The concentration curve controlled by the parameters estimated by the proposed method. The RMSE is 0.0334mg/ml (1.62% of c_m) in coarse sand media with a velocity of 0.10cm/min in case 1.

There is another question we should consider: whether the parameters estimated by the present method at observed point x_1=30cm could fit observed data at point x_2=70cm well or not. One could use the parameters estimated (n=0.13, u=0.565 cm/min, and D_L=0.083cm^2/min) to create theoretical BTC at observed point x_2=70cm,

and calculate RMSE using the observed data at x_2=70cm. A MATLAB script in Appendix B-3 could facilitate such a computational procedure which concludes that the RMSE at x_2=70cm based on the present method using observed data at x_1=30cm is 0.2369mg/ml.

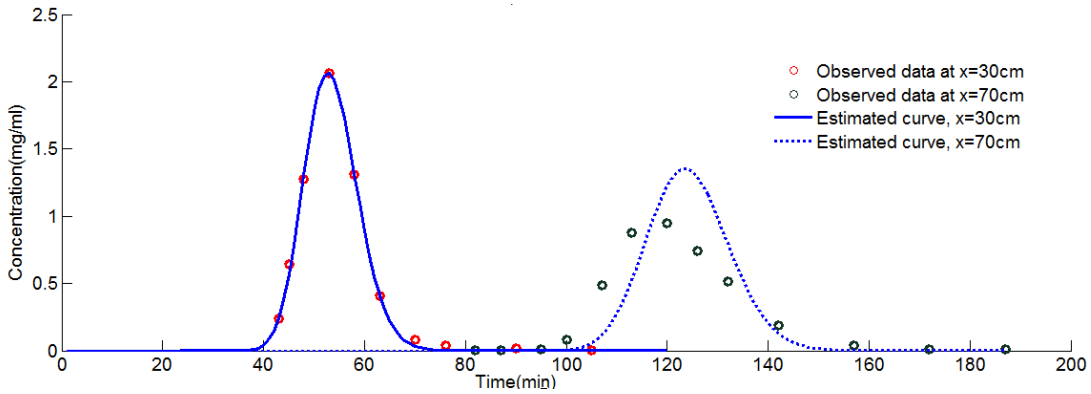


Figure 9. The concentration curves with the parameters n=0.13, u=0.565 cm/min, and D_L=0.083cm^2/min for observation points at x=30cm and 70cm. The RMSE is 0.0334mg/ml for estimation at x=30cm, and the RMSE is 0.2369 mg/ml for estimation at x=70cm in case 1.

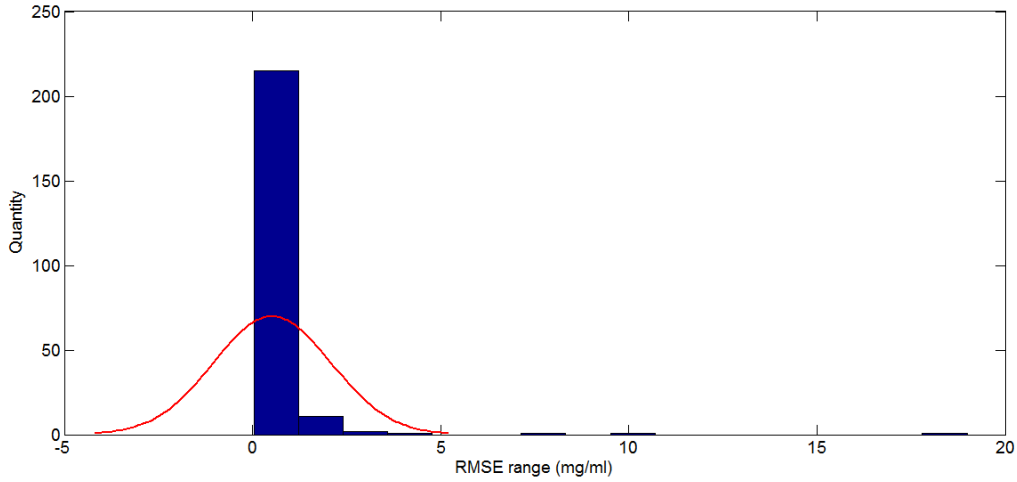


Figure 10. Histogram of all RMSE at x=70cm in case 1.

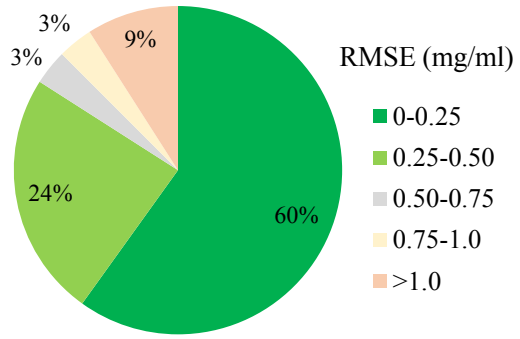


Figure 11. RMSE relative frequency distribution chart at x=70cm for case 1.

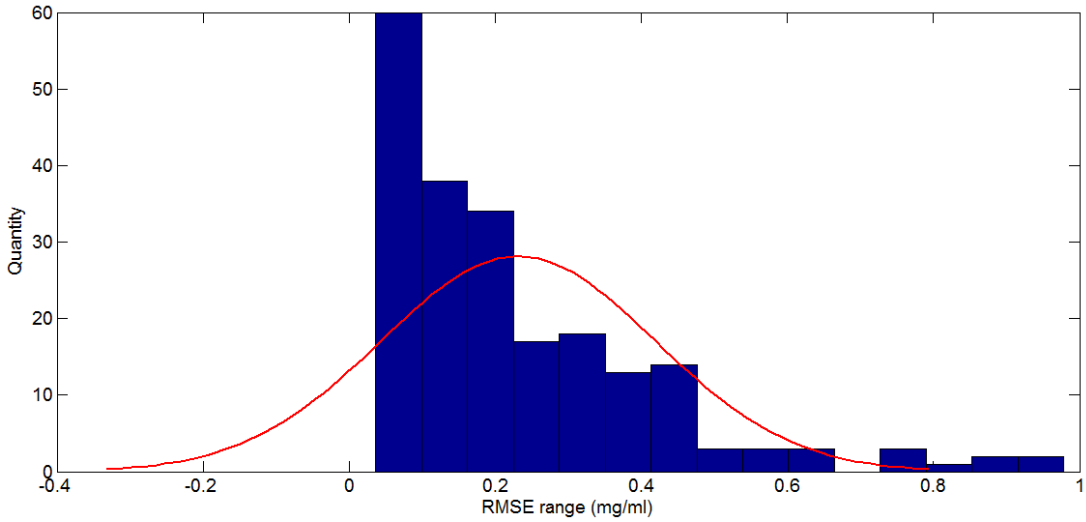


Figure 12. Histogram for RMSE<1mg/ml at x=70cm for case 1.

If selecting the observed point at x_2=70cm, picking three observed times and corresponding concentrations, now observed data at x_1=30cm would offer some references to check if the proposed method could estimate parameters fitting the observed point x_1=30cm satisfactorily.

The RMSE ranges from 0.0377mg/ml (3.99% of c_m=946mg/L) to 18.9861mg/ml, the minimum RMSE occurs when we choose t_1=107min

(c_1=484mg/L), t_2=120min (c_2=c_m=946mg/L), and t_3=132min (c_3=516mg/L), while in this case, the peak concentration c_m 946mg/L is tested at t_m=120min. So for this situation the optimal selection for the smallest RMSE requires that one chooses three observation times (at objective point x_2=70cm) as following steps: the first observation time (which is t_1) is the time for 50% of peak concentration to appear before the peak time the second observation time (which is t_2) is the peak time (t_m), and the third observation time (which is t_3) is the time for 50% of peak concentration to appear again after the peak time. On the contrary, the deviation from the actual data would be considerable if one selects all three observed data after peak time, because 21 combination sets of three times selected after peak time have a relatively high RMSE over 1mg/ml.

The RMSE relative frequency distribution shown by Figure 10, 11, and 12 illustrate that 60% of the total 232 combination selections creates 0-0.25mg/ml RMSE (26.43% c_m) using the parameters estimated by the proposed method, 24% of sets have the result of RMSE above 0.25mg/ml but below 0.50mg/ml (52.85%c_m), 9% of sets have result deviated significantly from the real data with a RMSE greater than 0.75mg/ml. The relative value of RMSE to the peak concentration at this observed point is twice as high as that at the observed point closer to inlet (x_1=30cm).

If we focus on the time sets which creates the smallest RMSE of 0.0377mg/ml (or 3.99% of c_m), the parameters could be estimated as: n=0.13, u=0.585 cm/min, and D_L=0.176 cm^2/min. If the effective molecular coefficient is neglected, then α_L=0.301cm. The Figure 13 below shows that BTCs at the observed point x_1=30cm with the parameters estimated by the proposed method has larger RMSE (which is

0.2371mg/ml) for data observed at x=30cm. In summary, the method used at observed point (x_2=70cm) far away from the inlet creates almost the same porosity and velocity as the one used at x_1=30cm. However, the estimated longitudinal dispersivity at x_2=70cm is twice larger than that at x_1=30cm.

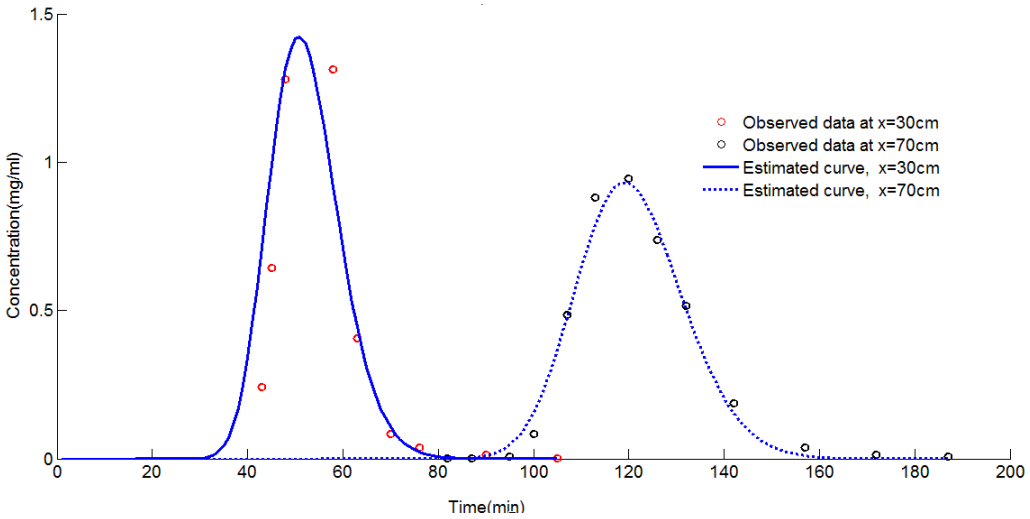


Figure 13. The concentration curves with the parameters n=0.13, u=0.585cm/min, and D_L=0.176cm^2/min at x=30cm and 70cm. The RMSE is 0.2371mg/ml for estimation at x=30cm. The RMSE is 0.0377 mg/ml for estimation at x=70cm with a velocity of 0.10cm/min in case 1.

If we use the Liang's method, both conservative and reactive solutes need at least two observed points for the 1-D case, which are assumed as x_1 and x_2. The peak times of the two observed point are t_{m1} and t_{m2}, respectively. Then the Liang's method estimates the parameters of velocity u and longitudinal dispersivity α_L as below,

$$u = \sqrt{\frac{t_{m2}x_1^2 - t_{m1}x_2^2}{t_{m1}t_{m2}(t_{m1}-t_{m2})}} \tag{91}$$

$$\alpha_L = \frac{1}{2}\left(\frac{x_1^2}{ut_{m1}} - ut_{m1}\right) \tag{92}$$

Here we could obtain x_1 =30cm, and x_2 =70cm, t_{m1} =53min, and t_{m2}=120min, respectively, then we could estimate,

$$u = 0.597 cm/min$$

This method provides the same velocity value as the proposed method.

As regarded to the CFM, we assume that the effective porosity estimated above is creditable, i.e. n=0.13. Based on the typical range of sand diameters of 0.5–0.7mm, and hydraulic conductivity K=3.72cm/min, we set the fitting range is from 0.05cm^2/min to 0.22cm^2/min with the interval as 0.002 cm^2/min for D_L, and 0.40cm/min to 0.80cm/min with the interval as 0.005cm/min for velocity u.

For the parameter ranges given above, the minimal RMSE could be obtained by running MATLAB script in Appendix B-4, which is 0.1404mg/ml. The maximal RMSE is 0.8357mg/ml. If one wants to obtain the minimal RMSE, i=36, j=4, which means that estimation for u and D should be 0.575cm/min, 0.106cm^2/min, respectively, and consequently the longitudinal dispersivity α_L is around 0.184cm.

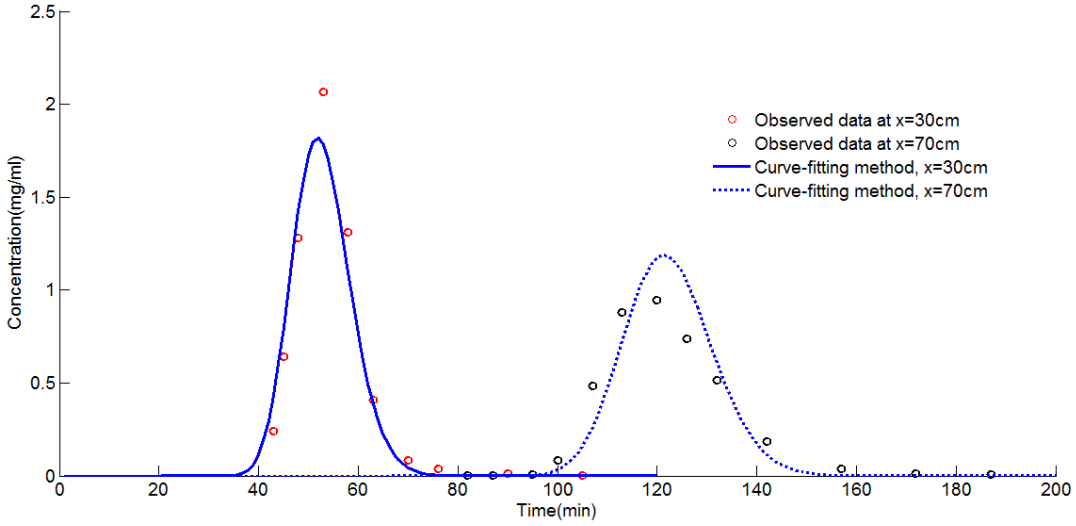


Figure 14. The concentration curves is created by curve-fitting method with the parameters as n=0.13, u=0.575 cm/min, and D_L=0.106 cm^2/min at x=30cm and 70cm. The RMSE is 0.1473mg/ml for estimation at x=30cm. The RMSE is 0.1348 mg/ml for estimation at x=70cm in case 1.

4.2 Estimation and comparison for case 2

Case 2: q=0.05 cm/min, sand diameters are around 0.5–0.7mm, K=3.72cm/min or 53.63m/d.

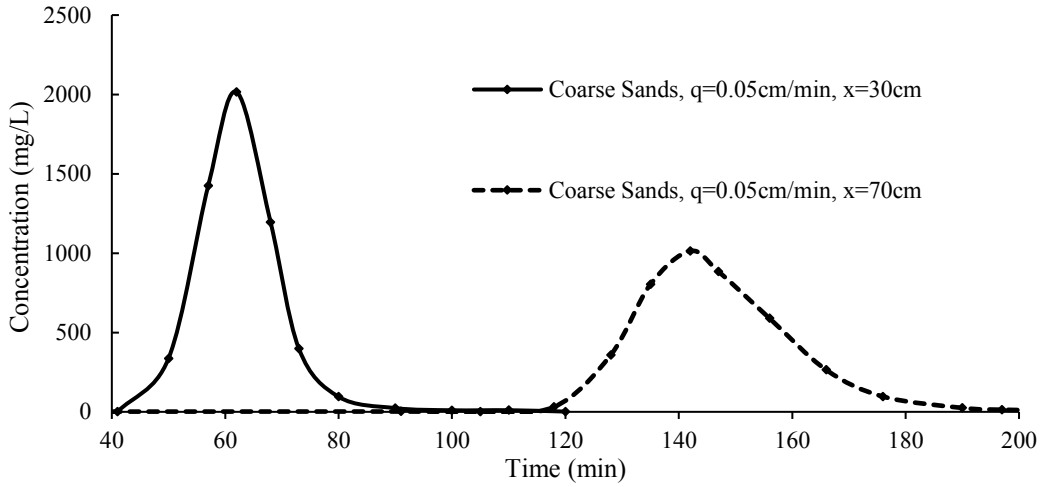


Figure 15. The breakthrough curves at two observation points in coarse sand column with a velocity of 0.05cm/min

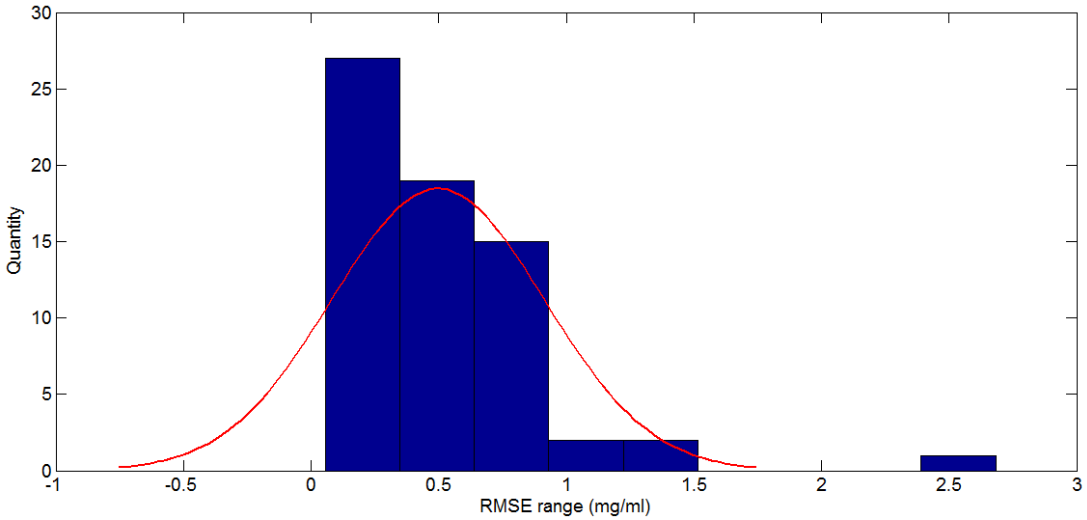


Figure 16. Histogram of all RMSE at x=30cm in case 2.

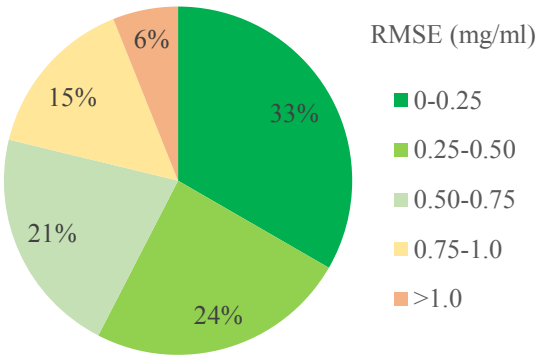


Figure 17. RMSE relative frequency distribution chart at x=30cm in case 2.

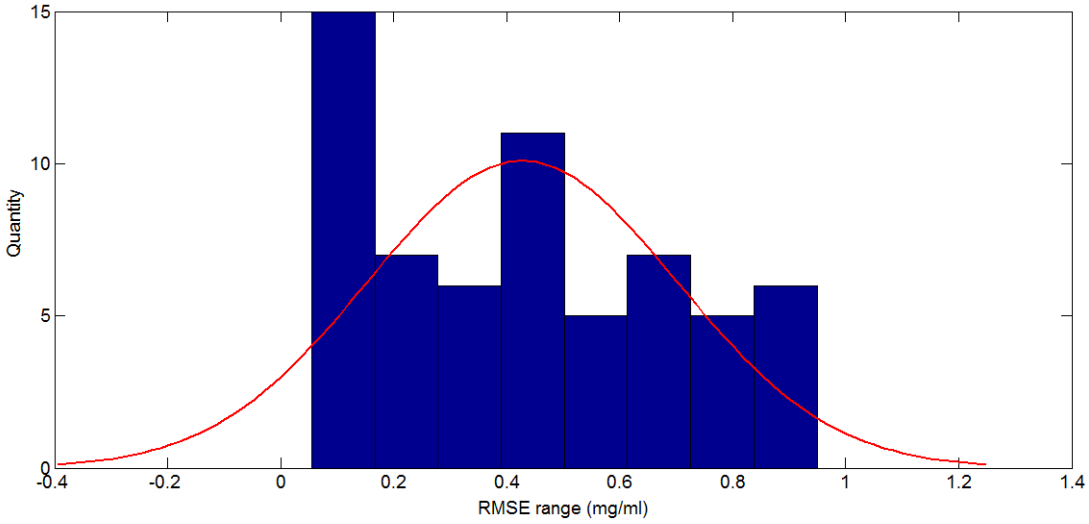


Figure 18. Histogram for RMSE<1mg/ml at x=30cm in case 2.

Firstly we focus on the observed point at x_1=30cm, one can pick three observed times and corresponding concentrations. Then one can use MATLAB script similar in

Appendix B-1 to obtain the final 66 combination sets of three different observed times, having different RMSE compared with observed data.

The RMSE ranges from 0.0550mg/ml (4.17% of c_m=1319mg/L) to 2.6832mg/ml, the minimum RMSE occurs when we choose t_1=34min (c_1=939mg/L), t_2=38min (c_2=c_m=1319mg/L), and t_3=42min (c_3=1096mg/L), while in this case, the peak concentration c_m 1319mg/L is tested at t_m=38min. So for this situation the best selection for the objective of the smallest RMSE needs one to choose three observation times (at objective point x_1=30cm) as follow the first observed time is the time point right before the peak time, the second observed time is the peak time, and the third observed time is the time point right after the peak time. On the contrary, if one chose t_1=38min (c_1=1319mg/L), t_2=46min (c_2=779mg/L), and t_3=56min (c_3=320mg/L), the deviation of the estimated BTCs from the actual data would be the greatest. This largest RMSE happens when we intensively select three observed times right after the peak time.

The RMSE relative frequency distribution could be analyzed by Figure 16, 17, and 18. As can be seen from the figures, 33% of the total 66 combination selection has the RMSE of 0-0.25mg/ml (which corresponds to 0%-18.95% of c_m) using the parameters estimated by the proposed method, 24% of sets has the result of RMSE above 0.25mg/ml but below 0.50mg/ml (which corresponds to 18.95%-37.91% of c_m), 21% of sets has the RMSE higher than 0.50mg/ml lower than 0.75mg/ml (which corresponds 37.91%-58.86% of c_m), the rest 21% of sets makes the result deviated more from the real data with the RMSE bigger than 1.00mg/ml.

Then if we focus on the time sets which create the smallest RMSE 0.0550mg/ml (4.17% of c_m), the parameters could be estimated as: n=0.1366, u=0.4840 cm/min, and D_L=0.066cm^2/min, if the effective molecular coefficient is neglected, α_L=0.135cm, as in Figure 19.

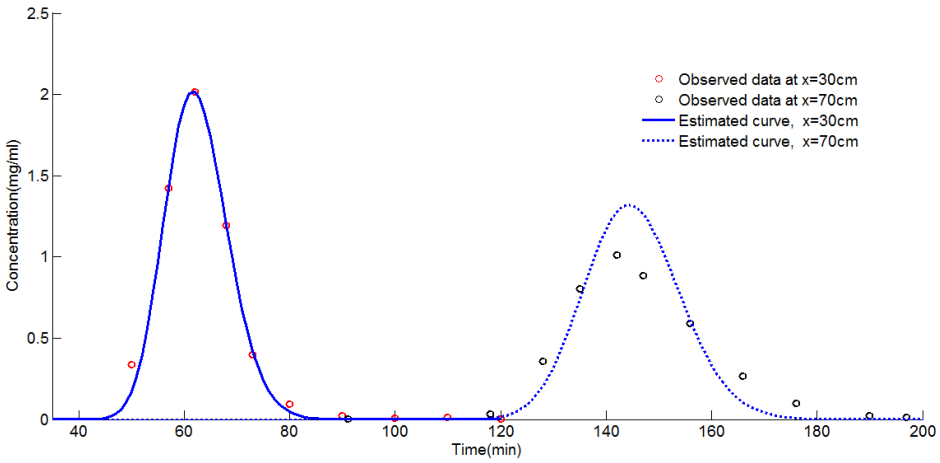


Figure 19. The concentration curve controlled by the parameters n=0.1366, u=0.4840 cm/min, D_L=0.066cm^2/min, and α_L=0.135cm estimated by the proposed method. The RMSE for observation point at x=30cm is 0.0550mg/ml. The RMSE for observation point at x=70cm is 0.1276mg/ml in case 2.

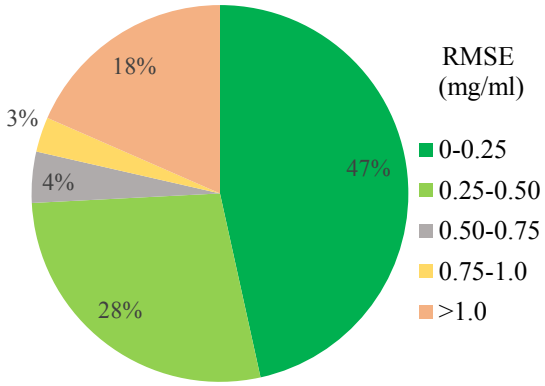


Figure 20. RMSE relative frequency distribution chart at x=70cm in case 2.

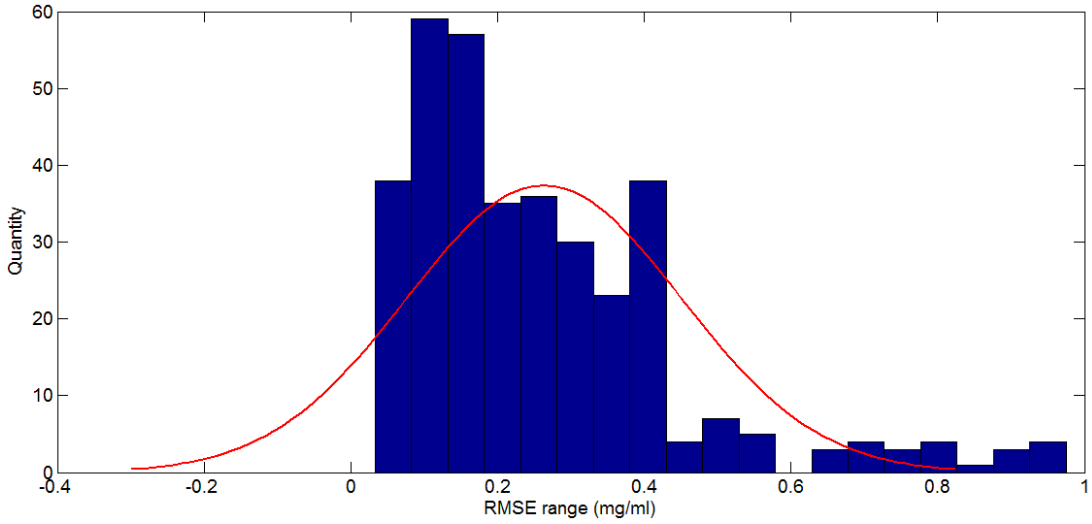


Figure 21. Histogram for RMSE<1mg/ml at x=70cm in case 2.

One could use the parameters estimated (n=0.1366, u=0.4830 cm/min, and D_L=0.066cm^2/min) to create theoretical BTCs at observed point x_2=70cm, and calculate the RMSE using the observed data at x_2=70cm. Figure 19 helps one to conclude the RMSE for x_2=70cm based on the present method using observed data at x_1=30cm is 0.1276mg/ml (12.60 % of c_m=1013mg/L).

Now if one selects the observed point at x_2=70cm, picks three observed times and corresponding concentrations, now observed data at x_1=30cm would offer reference standard to check if the proposed method could estimate parameters fitting the data at observed point x_1=30cm well.

The RMSE ranges from 0.0333mg/ml (3.29% of c_m=1013mg/L) to over 10mg/ml, the minimum RMSE occurs when we choose t_1=135min (c_1=803mg/L), t_2=142min (c_2=c_m=1013mg/L), and t_3=156min (c_3=589mg/L), while in this case, the

peak concentration c_m 1013mg/L is tested at t_m=142min. So for this situation the best selection for the objective of the smallest RMSE needs one to choose three observed time (at objective point x_2=70cm) as follow the first one is the time point right before peak time, the second one is the peak time, and the third one is the time point after the peak time but close to the peak time. On the contrary, the deviation from the actual data would be greater if one selects all three observed times at the very beginning of the increasing limb, or at the very end of the decreasing limb.

Figure 19, 20, and 21could be used to explain the RMSE relative frequency distribution. This figure illustrates that 47% of the total 434 combination selection has the RMSE ranging from 0 to 0.25mg/ml (which corresponds to 0%-24.68% of c_m) using the parameters estimated by the proposed method, 28% of sets has the RMSE above 0.25mg/ml but below 0.50mg/ml (which corresponds to 24.68%-49.36% of c_m), 25% of sets presents the result deviated more from the real data with the RMSE bigger than 0.75mg/ml.

If we focus on the time sets which has the smallest RMSE 0.0333mg/ml (3.29% of c_m), the parameters could be estimated as: n=0.13, u=0.4879 cm/min, and D_L=0.1192 cm^2/min, if the effective molecular coefficient is neglected, α_L=0.2443cm, as in Figure 22 below.

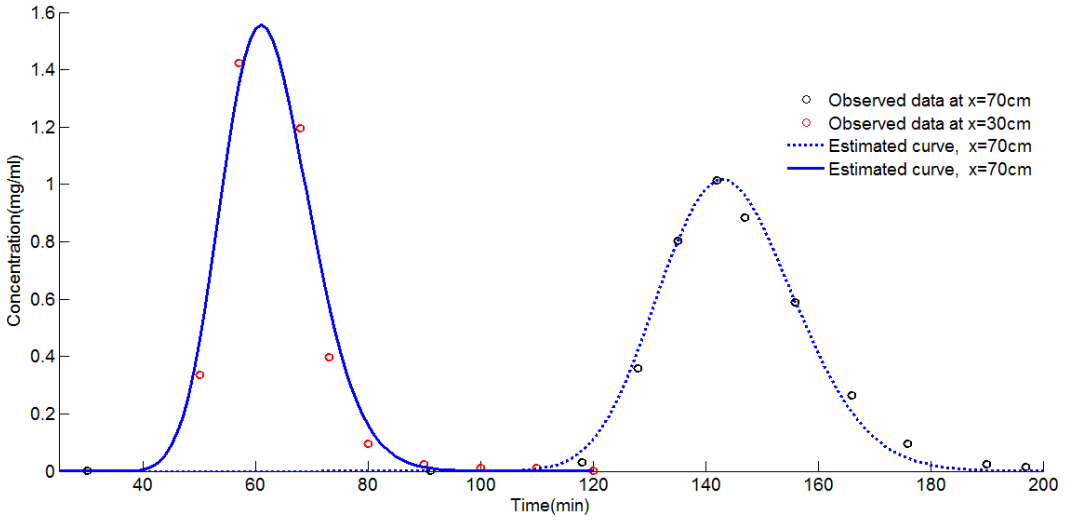


Figure 22. The concentration curves with the parameters n=0.13, u=0.488 cm/min, and D_L=0.1192 cm^2/min at x=30cm and 70cm. The RMSE is 0.0883mg/ml at x=30cm. The RMSE is 0.0333 mg/ml at x=70cm in case 2.

One could use the parameters estimated (n=0.13, u=0.488 cm/min, and D_L=0.1192cm^2/min) to create theoretical breakthrough curve at observed point x_1=30cm, and calculate RMSE using the observed data at x_1=30cm. Figure 22 illustrates the RMSE for x_1=30cm based on the present method using observed data at x_2=70cm is 0.0883mg/ml (12.60 % of c_m=1013mg/L). Overall, the low value of the RMSE means the estimated BTCs fit well with the observed data which proves that the proposed method provides credible parameter estimation result.

If we use the Liang's method, the peak times of the two observed points are t_{m1} and t_{m2}, respectively. Here we could obtain x_1 =30cm, and x_2 =70cm, t_{m1} =62min, and t_{m2}=142min, respectively, then we could use Eq. (91) to estimate

$$u = 0.500 cm/min$$

This method provides almost the same velocity value as the proposed method.

Considering the CFM, we assume the porosity measured in laboratory is correct, n=0.13. Based on the normal value range for sand whose diameters are around 0.5–0.7mm, hydraulic conductivity K=3.72cm/min, we set the fitting range is from 0.10cm^2/min to 0.22cm^2/min with the interval as 0.002 cm^2/min for D_L, and 0.40cm/min to 0.80cm/min with the interval as 0.005cm/min for velocity u.

For the parameter ranges given above, the minimal RMSE could be obtained by running MATLAB script in Appendix B-4, which is 0.0613mg/ml. If one wants to obtain the minimal RMSE, estimation for u and D would be 0.485cm/min, 0.100cm^2/min, then the longitudinal dispersivity α_L is around 0.206cm.

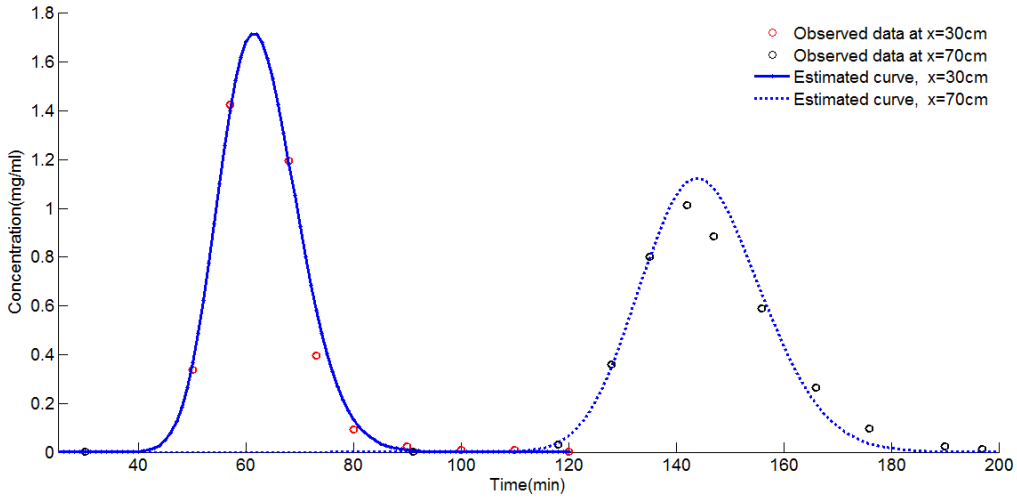


Figure 23. The BTCs is created by the CFM with the parameters n=0.13, u=0.485 cm/min, and D_L=0.100 cm^2/min at x=30cm and 70cm. The RMSE is 0.0643mg/ml at x=30cm. The RMSE is 0.0597 mg/ml at x=70cm in case 2.

Table 4. Summary of estimated parameters using proposed method in two cases.

Case	Observation point	Estimated D_L (cm^2/min)	Estimated Velocity u (cm/min)	Estimated α (cm)	RMSE (mg/L)
Case-1 K=3.72cm/min n=0.32 q=0.10cm/min	x=30cm	0.0830	0.5650	0.1469	180
	x=70cm	0.1760	0.5850	0.3009	216
Case-2 K=3.72cm/min n=0.32 q=0.05cm/min	x=30cm	0.0660	0.4840	0.1364	107
	x=70cm	0.1192	0.4880	0.2443	59

Table 5. Summary of estimated parameters using curve-fitting and Liang's method.

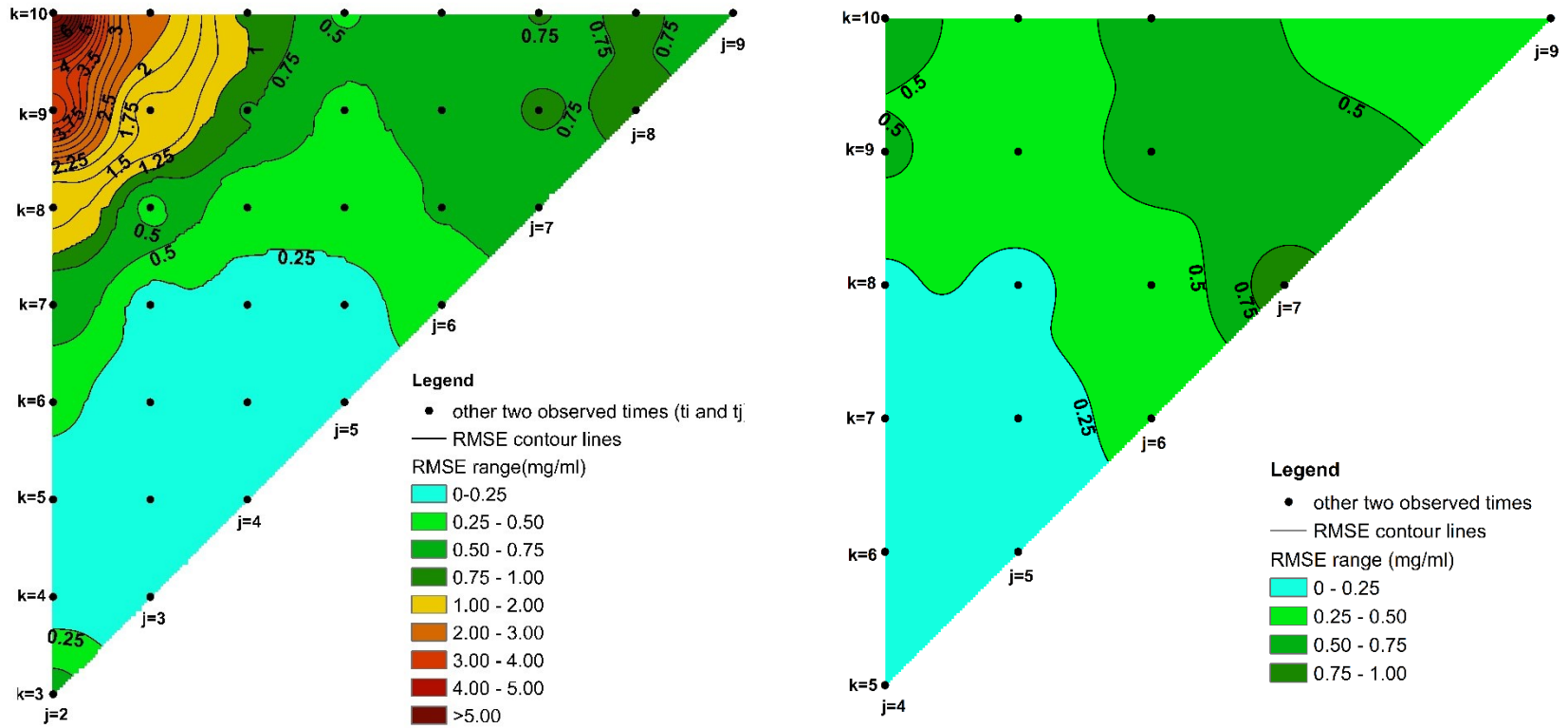
Case		Estimated D_L (cm^2/min)	Estimated Velocity u (cm/min)	Estimated α (cm)	RMSE (mg/L)		Estimated Velocity (cm/min)	Estimated α (cm)
Case-1 K=3.72cm/min n=0.32 q=0.10cm/min	Curve-fitting method	0.1060	0.5750	0.1843	140	Liang's method	0.5970	0.1950cm coarse sands column [Liang et al. (2018)]
Case-2 K=3.72cm/min n=0.32 q=0.05cm/min		0.1000	0.4850	0.2062	61		0.5000	

4.3 Uncertainty analysis

As discussed case by case, the main error of this proposed method for interpreting the BTCs and parameter estimation is associated with the selection of observed times, which is a key factor in uncertainty analysis. Now we only focus on the generalized model for 1-D conservative instantaneous source test, one should choose three sets of sample data (at the same observed point) to estimate three unknown parameters (u, n, α_L). If there are N sets of data $((t_1, C_1), (t_2, C_2),\ldots,(t_N, C_N))$ collected at a certain observed point in the laboratory column test, we should consider

cases of total C_N^3 which will lead to different interpretation and estimation result. In each case selected three pairs of observed times and corresponding concentrations could be generally recorded as (t_i, C_i), (t_j, C_j), and (t_k, C_k), $1 \leq i < j < k \leq N$ is required. The number of i (or the first selected data point) plays a leading role in determining how these three pairs of data located on BTCs, which could indicate the fore-and-aft distribution comparing to the peak time (t_m).

So here for the purpose of uncertainty analysis, I firstly settle down the specific number value i (the first selected data point), summarize the RMSE from different cases with different number values of j and k (or the second and third selected data points), which means j and k could consist a plane with the orthogonal coordinates and each coordinate value is assigned by RMSE. To evaluate the interpretation on this jk plane, one could utilize ArcGIS for result visualization. The principle tool I used to create Figure 24, 25, 26, and 27 is an inverse distance weighted (IDW) technique and Contour in Spatial Analyst.



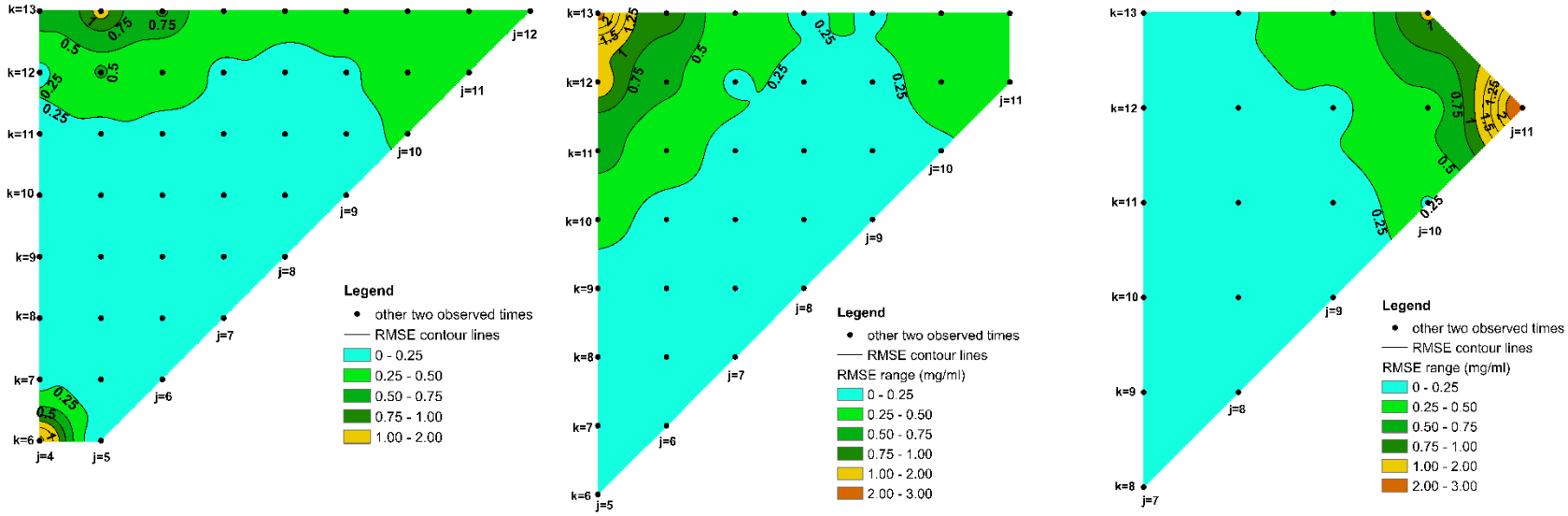
A B

Figure 24. In Case 1 and the observation point is x=30cm.The number of the first observed time is i, which is also the order number for the column of time in Appendix A-1. Peak time (t_m) is estimated at Number m=6 in the total 13 observed times. i=1 for the left figure (A) and i=3 for the right figure (B). $1 \leq i < j < k \leq 13$ is required. The interval is set as 0.25mg/ml for RMSE contour lines, and the break value is unified to create the Color Scale Bar. The black point refers to the set of three pairs of observed times.

All the information is shown in right triangles which are restricted upper left above the imaginary line $j=k$. Sector distribution is obvious in both charts in Figure 24, with the smallest RMSE value centered in the left bottom, peak time m=5 is included. It means when the two observed times are selected closely to each other surrounding near the peak time, the interpretation quality is ideal, resulting the RMSE less than the value of 0.25mg/ml. Inversely, the top right corner centralizes the bigger RMSE value. It implies the estimation error would be increased if the other two observed times are selected very late, at the end of the decreasing limb in BTCs.

For i=1 in Figure 25-A, the contour lines are more dense at the top left corner, and the range of RMSE values is enlarged mainly by those points. It implies the case when one picks the observed time very early and really late, at the beginning of the increasing limb and at the end of decreasing limb, respectively, would lead to increasing the estimation uncertainty. Such uncertainty could be alleviated in Figure 24-B when the first observed time is picked a little bit later (i=3).

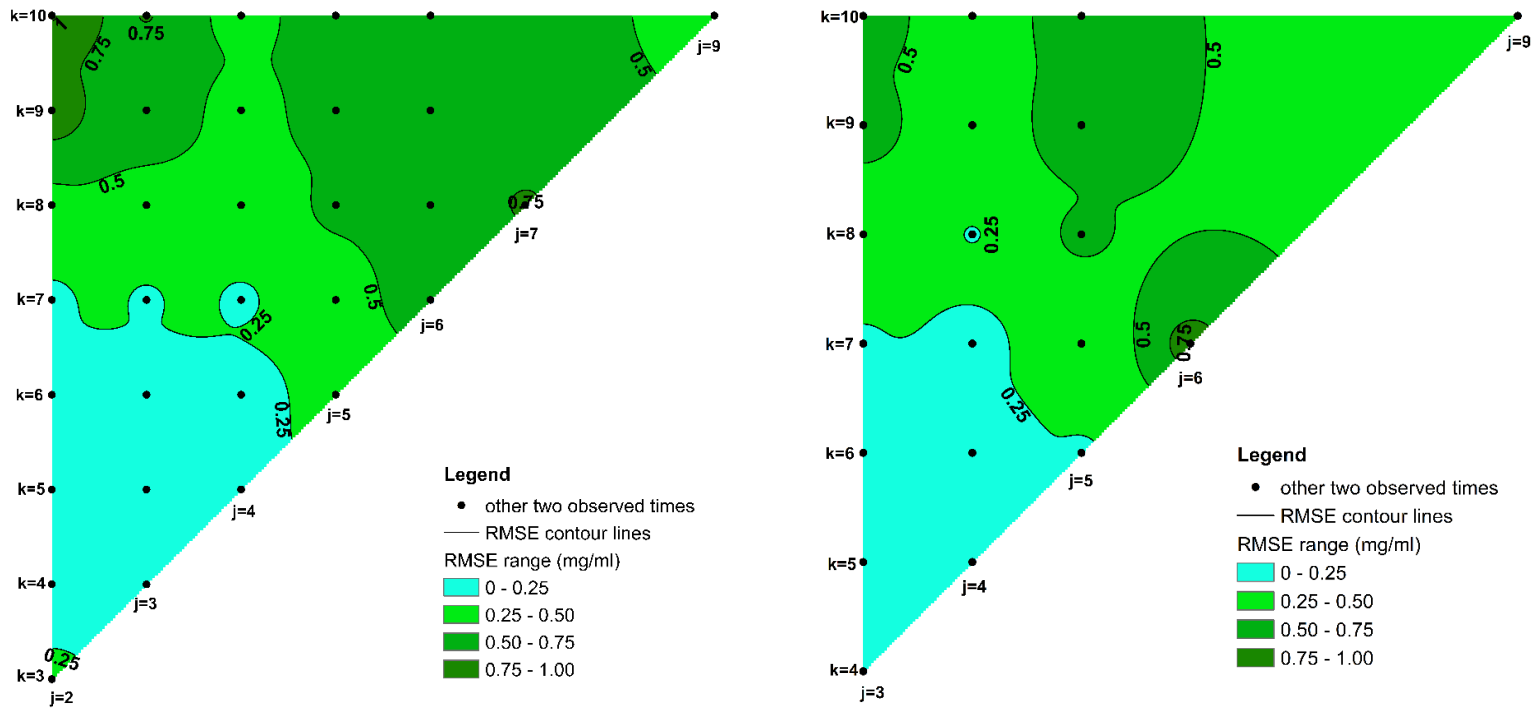
In case 1, at x=70cm, I assign the first observed time as Number value i=1, i=4, and i=6, which is more and more closing to the peak time (Number m=8), and obtain the Figure 25-A, 25-B, and 25-C, respectively. If we compare Figure 24-A with Figure 25-A (i is both equal to 1) or Figure 24-B with Figure 25-B (with the similar value of i), uncertainty of parameters estimation is alleviated because the belt informing the lower value of RMSE occupies larger portion of area in Figure 25 (x=70cm) than that in Figure 24 (x=30cm). The dispersion would play a more significant role for BTCs theoretically at observation points with longer distances, so these observation points would provide better estimates of parameter values with less uncertainty.



A B C

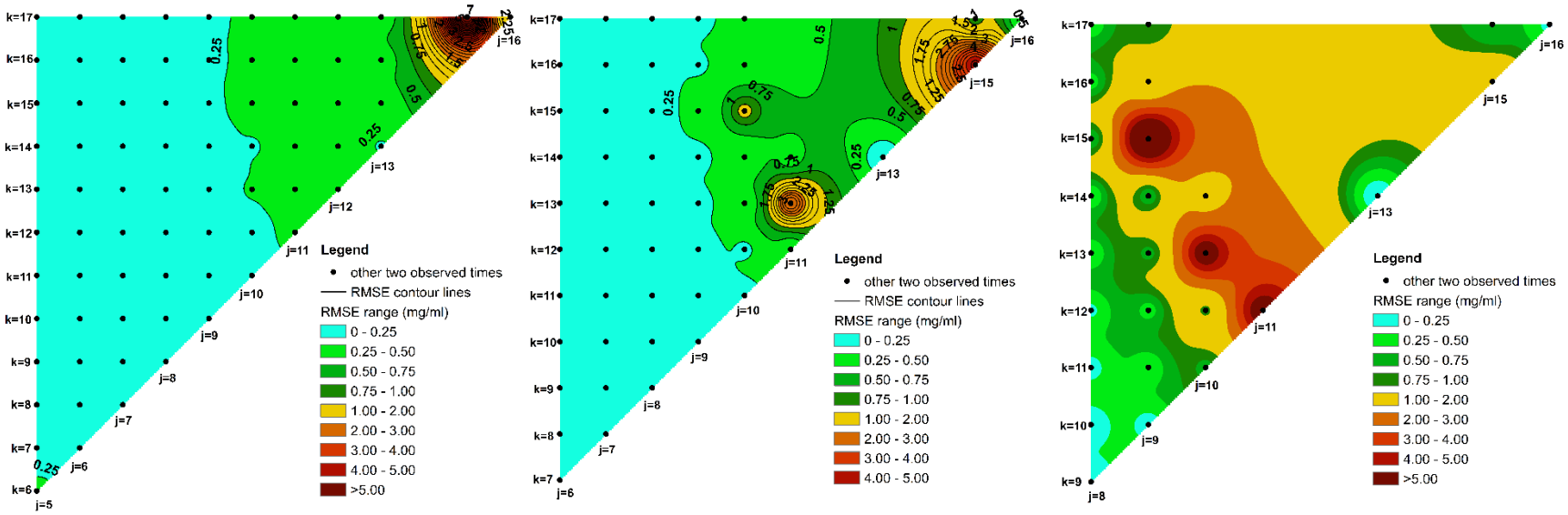
Figure 25. In Case 1 and the observation point is x=70cm.The number of the first observed time is i, which is also the order number for the column of time in Appendix A-1. Peak time (t_m) is estimated at Number m=8 in the total 16 observed times. i=1 for the left figure (A), i=4 for the middle figure (B), and i=6 for the right figure (C). $1 \leq i < j < k \leq 16$ is required. The interval is set as 0.25mg/ml for RMSE contour lines, and the break value is unified to create the Color Scale Bar. The black point refers to the set of three pairs of observed times.

In the domain with a lower uniform velocity, like in case 2, decreasing from 0.10cm/min to 0.05cm/min, the distance of the observation point from the instantaneous source would take more significant role in increasing the uncertainty for interpreting the BTCs and estimating parameters. In the domain with lower uniform velocity, like in case 2, decreasing from 0.10cm/min to 0.05cm/min, the distance of the observation point from the instantaneous source would take more significant role in the uncertainty for interpreting the BTCs and estimating parameters. For the observation point x=30cm, Figure 26 shows that the sector distribution of the value of RMSE is overwhelmingly restricted under the level of 1.00mg/ml, and the best estimation would be achieved when the three observed times are all near the peak time (m=4), but it does not mean the peak time must be selected, the observed times which are in the range of two time-interval difference could be used and combined to estimate the parameters with less uncertainty. For the observation point x=70cm, the uncertainty would increase significantly when the practitioner only observed the times all at the decreasing limb in BTCs. The nearly vertical belt with lowest value of RMSE in color light blue is both observed in Figure 27-A (i=2) and Figure 27-B (i=5), it implies that the earlier observation time is required if one wants to interpret the BTCs (j has lower value at left side). However at x=70cm, the peak time is observed at m=5th time point, so when we determine the first observation time much later after the peak time, like in Figure 27-C i is set as 7 (two time-interval difference), the estimation would not be creditable and the uncertainty for interpreting the BTCs would increase.



A B

Figure 26. In Case 2 with the lower uniform velocity and the observation point is x=30cm.The number of the first observed time is i, which is also the order number for the column of time in Appendix A-2. Peak time (t_m) is estimated at Number m=4 in the total 11 observed times. i=1 for the left figure (A) and i=2 for the right figure (B). $1 \leq i < j < k \leq 11$ is required. The interval is set as 0.25mg/ml for RMSE contour lines, and the break value is unified to create the Color Scale Bar. The black point refers to the set of three pairs of observed times.



A B C

Figure 27. In Case 2 with the lower uniform velocity and the observation point is x=70cm.The number of the first observed time is i, which is also the order number for the column of time in Appendix A-2. Peak time (t_m) is estimated at Number m=5 in the total 16 observed times. i=2 for the left figure (A), i=5 for the middle figure (B), and i=7 for the right figure (C). $1 \leq i < j < k \leq 16$ is required. The interval is set as 0.25mg/ml for RMSE contour lines, and the break value is unified to create the Color Scale Bar. The black point refers to the set of three pairs of observed times

According to the uncertainty analysis discussed above, the recommendations could be summarized as follows if the proposed method is applied into practice. First, the proposed method appears to be a robust and creditable assessment tool applicable for estimating parameters with acceptable estimation errors. Second, among the three selected data points for the 1D case, as long as the concentration of the second data point is larger than those of the first and third ones, the method is reliable for parameter estimation. Third, if the concentrations of all three data points are monotonically decreasing (i.e., such three points are all in the declining limb of the BTCs), the method becomes less reliable.

4.4 Application and limitations

In summary, this new method unlike the Liang method (Liang et al., 2018) whose accuracy relies heavily on the peak time determination, is not subsequently controlled by the time-interval of measurement. The error or the uncertainty of the parameter estimation can be substantially reduced if one can determine three pairs of observed time and concentration at certain observation point surrounding the peak time. Before doing so, one has to have a rough approximation of the peak time**.**

This can be done using the following steps. Firstly, based on at least three spatial head observation points, one can determine the flow direction with low cost and then carry out a test with a relatively large time interval of measurement to estimate the pore velocity that is much less sensitive to the time-interval of measurement, as demonstrated earlier. Secondly, one can calculate the advective time based on the estimated pore velocity and the location of the observation point. It is notable that the peak time is somewhat less than the advective time due to the

dispersive effect. Finally, one can refine the time interval of measurement before the advective time but not be required to accurately determine the peak time. The same procedure can be used for the 2-D and 3-D cases as well.

Within sufficient financial support and time (which is sometimes not the case from a practical stand of point), the irreplaceability of the programmed parameter estimation methods, such as CXTFIT, UCODE, and PEST should be highly acknowledged. However the flexibility and easy implementation of the proposed analytical method are attractive to scientists and practitioners in analyzing subsurface transport data in practice. For better complementary application, we should address some notable issues of this study. Installation of multiple sensors at different distances from the entrance of a long sand column is a quite common practice. However, it is different for some laboratory column experiments in which only a single observation point is available (usually at the end of the column). As mentioned above, the advantage of the proposed method is it contains the solution for the case with only one observation point. The proposed method is based on the analytical solutions for ADE in 1-D, 2-D and 3-D spaces. Therefore, some simplifications of porous media, flow field, and boundary conditions are necessary. However, it cannot capture the details of transport in a heterogeneous porous media, for which numerical simulations are somehow more adequate, so developing flexible and simple parameter estimation methods for transport in heterogeneous field sites should be pursued in the future.

5. CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

This study proposes a simple method to estimate the transport parameters (the pore velocity, dispersivity, and porosity) of conservative and reactive solutes in 1-D, 2-D and 3-D domains with uniform flow fields based on the measured BTCs of the instantaneous tracer tests. The reactive solutes transport is limited to the linear sorption isotherm and the first-order decay. The main advantages of this method are: 1) It requires fewer measured data than the traditional curve-fitting methods, three pairs of observation data for 1-D domain, four for 2-D domain, and five for 3-D domain; 2) It does not require any specific optimized operation as the parameter estimation is done in closed-form algebraic functions. The proposed method is applied on tracer tests in the laboratory sand columns. The estimated parameters are close to the CFM, the estimated velocity is close to the Liang's method. The main error of the parameter estimates is associated with the selection of three time points, which is a key factor for the proposed method of this study. The error of the estimated pore velocity is very small compared to the Liang's method and curve-fitting method. While the error of the estimated dispersivity compared to curve-fitting method increases when the velocity is slow.

This method used four case to prove that the three observation times selection is important to the estimation of parameters, if one pick the time on the left part in increasing tail (before the peak time), peak time, and the time on the right part in the decreasing part (after the peak time), the error of estimated parameters could be small, however, one should avoid to choose the three time points all on the one tailing side

or at very beginning and very end of the curve. In sum this method can be employed easily by scientists and practitioners for parameter estimations in laboratory column experiments if advection-dispersion equation is applicable. As the governing equation of heat transport is essentially the same as contaminant transport in the subsurface, provided that heat transport parameters are employed instead of the contaminant transport parameters, the method developed in this study can also be used for parameter estimation of heat transport in a laboratory column test if a slug heat source (instead of a slug tracer source) is injected in porous media with the presence of a uniform flow field.

5.2 Future work

There are a number of notable issues for better applications of this study. First, even though theoretically the proposed method only needs three observation time points at one observation location to obtain the parameters estimation, requiring fewer measured data at one certain observation point than the traditional CFM using all observation points, the selection of observation times should not be very casual for the laboratory column test. As mentioned above, if one select all observed times at the very beginning of the increasing limb or at the very end of the decreasing limb of BTCs the parameters estimation could be both incredible. Best estimation with the lowest RMSE value generally requires one to pick the first observed time before but close to the peak time, the second observed time just at the peak time, and the third observed time after bur also close to the peak time. The latent requirement is that practitioners are expected to obtain the BTCs before the time selection, at least the period around the peak time. However, the flexibility and easy implementation of the

proposed method are attractive to scientists and practitioners in analyzing subsurface transport data in practice.

Second, the proposed method is based on the analytical solutions for ADE in 1-D, 2-D and 3-D spaces. Therefore, some simplifications have been employed for porous media, flow field, and boundary conditions. For instance, the transport parameters such as dispersion coefficients and flow velocities are assumed to representatives of a "homogenized" porous media. ADE is found to be problematic in some cases, particularly for transport in highly heterogeneous porous media, where the contaminant transport requires more complex and integrated analytical solution. For the simplified and generalized model used in Chapter 1, experimental data are still lacked to test if the proposed method could estimate credible parameters.

Third, instantaneous injection assumption will become a problem when applying the proposed method on some field sites. Even though the continuous injections with a constant concentration is already simplified for various applications in the real world, the corresponding solution should integrate the injecting time to the analytical equation.

REFERENCES

Jacob Bear (1972). Dynamics of Fluids in Porous Media. Environmental Science Series (New York, 1972). xvii. American Elsevier Pub. Co., New York, 764.

Jacob Bear (1961). Some experiments in dispersion. *Journal of Geophysical Research*, 66 (8):2455–2467.

D.M. Mackay, D.L. Freyberg, P.V. Roberts (1986). A natural gradient experiment on solute transport in a sand aquifer. 1. Approach and overview of plume movement. *Water Resources Research*, 22 (13), 2017–2029.

Jie Ma, Huaming Guo, Mei Lei, Yongtao Li, Liping Weng, Yali Chen, Yuling Ma, Yingxuan Deng, Xiaojuan Feng, Wei Xiu (2018). Enhanced transport of ferrihydrite colloid by chain-shaped humic acid colloid in saturated porous media. *Science of the Total Environment*, 621, 1581–1590.

M.Boy-Roura, J.Mas-Pla, M.Petrovic, M.Gros, D.Soler, D.Brusi, A.Menció (2018). Towards the understanding of antibiotic occurrence and transport in groundwater: findings from the Baix Fluvia alluvial aquifer (NE Catalonia, Spain). *Science of the Total Environment*, 612, 1387–1406.

Bing Han, Wen Liu, Xiao Zhao, Zhengqing Cai, Dongye Zhao (2017). Transport of multi-walled carbon nanotubes stabilized by carboxymethyl cellulose and starch in saturated porous media: influences of electrolyte, clay and humic acid. *Science of the Total Environment*, 599, 188–197.

Lei Wu, Huanyu Chang, Xiaoyi Ma(2017). A modified method for pesticide transport and fate in subsurface environment of a winter wheat field of Yangling, China. *Science of the Total Environment*, 609, 385–395.

Xueyan Lv, Bin Gao, Yuanyuan Sun, Shunan Dong, Jichun Wu, Beilei Jiang, Xiaoqing Shi (2016). Effects of grain size and structural heterogeneity on the transport and retention of nano-TiO_2 in saturated porous media. *Science of the Total Environment*, 563, 987–995.

L. Weaver, N. Karki, M. Mackenzie, L. Sinton, D. Wood, M. Flintoft, P. Havelaar, M. Close (2016). Microbial transport into groundwater from irrigation: comparison of two irrigation practices in New Zealand. *Science of the Total Environment*, 543, 83–94.

Xiuyu Liang, Hongbin Zhan, Jin Liu, Guiming Dong, You-Kuan Zhang (2018). A simple method of transport parameter estimation for slug injecting tracer tests in porous media. *Science of the Total Environment*, 644, 1536-1546.

Shlomo P. Neuman, Daniel M. Tartakovsky (2009). Perspective on theories of non-Fickian transport in heterogeneous media. *Advances in Water Resources*, 32 (5), 670–680.

G. De Josselin De Jong (1958). Longitudinal and transverse diffusion in granular deposits. *Eos Transactions American Geophysical Union*, 39 (1), 67–74.

P. A. Domenico, F. W. Schwartz (1990). Physical and Chemical Hydrogeology. xxii. Wiley, New York, 824.

Jean-Pierre Sauty (1980). An analysis of hydrodispersive transfer in aquifers. *Water Resources Research*, 16(1), 145–158.

Eungyu Park, Hongbin Zhan (2001). Analytical solutions of contaminant transport from finite one-, two-, and three-dimensional sources in a finite-thickness aquifer. *Journal of Contaminant Hydrology,* 53, 41-61.

J. C. Parker, M. Th. Van Genuchten (1984). Determining transport parameters from laboratory and field tracer experiments. *Virginia Agricultural Experiment Station Bulletin*, 84–3, 1–96.

N. Toride, F. J. Leij, M. Th. van Genuchten (1995). The CXTFIT Code for Estimating Transport Parameters From Laboratory or Field Tracer Experiments, 2. Version 2, Research Report. No. 137. U. S. Salinity Laboratory, USDA, ARS, Riverside, CA.

L. Li, D. A. Barry, J. Morris, F. Stagnitti (1999). CXTANNEAL: an improved program for estimating solute transport parameters. *Environmental Modelling & Software*, 14 (6), 607–611.

E. P. Poeter, M. C. Hill (1999). UCODE, a computer code for universal inverse modeling. *Computers & Geosciences*, 25 (4), 457–462.

J. Doherty (2001). PEST-ASP Users' Manual. Watermark Numerical Computing, Brisbane, Australia.

E. P. Poeter, M. C. Hill, D. Lu, C. R. Tiedeman, S. Mehl (2014). UCODE_2014, With New Capabilities to Define Parameters Unique to Predictions, Calculate Weights Using Simulated Values, Estimate Parameters With SVD, Evaluate Uncertainty With MCMC, and More. *Integrated Groundwater Modeling Center Report,* GWMI 2014-02.

APPENDIX A: Laboratory data

A-1. Laboratory data for case 1. Coarse sand column with a velocity of 0.10cm/min

x=30cm, N=13 $G_{TRef} = 2.9549*c(Cl^-) + 950.83$				
Time (min)	Conductivity (µS/cm)	Temperature (°C)	25°C Conductivity (µS/cm)	Total Cl^- (mg/L)
23	724	13.5	940.260	0.000
35	726	13.5	942.857	0.000
43	1314	14.5	1663.291	241.112
45	2277	14.9	2853.383	643.864
48	3830	15.5	4728.395	1278.407
53	5800	16.1	7055.961	2066.104
58	4000	16.4	4830.918	1313.103
63	1810	17.0	2154.762	407.436
70	1037	18.2	1200.231	84.403
76	911	17.9	1061.772	37.545
90	855	18.0	994.186	14.673
105	836	18.8	954.338	1.187
124	793	18.1	919.954	0.000
x=70cm, N=16 $G_{TRef} = 2.9549*c(Cl^-) + 950.83$				
Time (min)	Conductivity (µS/cm)	Temperature (°C)	25°C Conductivity (µS/cm)	Total Cl^- (mg/L)
75	822	18.7	940.503	0.000

A-1 continued

x=70cm, N=16 G_{TRef} = 2.9549*c(Cl^-) + 950.83				
Time (min)	Conductivity (µS/cm)	Temperature (℃)	25℃ Conductivity (µS/cm)	Total Cl^- (mg/L)
82	820	18.1	951.276	0.151
87	819	18.0	952.326	0.506
95	866	19.3	977.427	9.001
100	1058	19.2	1196.833	83.252
107	2086	18.8	2381.279	484.094
113	3111	18.8	3551.370	880.077
120	3252	18.4	3746.544	946.128
126	2689	17.9	3134.033	738.841
132	2149	18.4	2475.806	516.084
142	1311	18.7	1500.000	185.851
157	919	18.1	1066.125	39.018
172	858	18.4	988.479	12.741
187	852	18.6	977.064	8.878
207	930	19.0	1056.818	0.000
217	830	19.0	943.182	0.000

A-2. Laboratory data for case 2. Coarse sand column with a velocity of 0.05cm/min

x=30cm, N=11 $G_{TRef} = 2.9549*c(Cl^-) + 950.83$				
Time (min)	Conductivity (µS/cm)	Temperature (℃)	25℃ Conductivity (µS/cm)	Total Cl^- (mg/L)
41	739	14.1	945.013	0.000
50	1524	14.2	1943.878	336.068
57	4056	14.3	5160.305	1424.575
62	5456	14.5	6906.329	2015.466
68	3541	14.5	4482.278	1195.116
73	1679	14.5	2125.316	397.471
80	975	14.6	1231.061	94.836
90	806	14.5	1020.253	23.494
100	776	14.6	979.798	9.803
110	775	14.5	981.013	10.214
120	757	14.6	955.808	1.685
x=70cm, N=16 $G_{TRef} = 2.9549*c(Cl^-) + 950.83$				
Time (min)	Conductivity (µS/cm)	Temperature (℃)	25℃ Conductivity (µS/cm)	Total Cl^- (mg/L)
105	747	14.5	945.570	0.000
118	828	14.7	1042.821	31.132
128	1576	14.2	2010.204	358.514
135	2606	14.2	3323.980	803.123
142	3085	14.1	3945.013	1013.294
147	2793	14.2	3562.500	883.844
156	2132	14.6	2691.919	589.221

A-2 continued

x=70cm, N=16 $G_{TRef} = 2.9549*c(Cl^-) + 950.83$				
Time (min)	Conductivity (µS/cm)	Temperature (℃)	25℃ Conductivity (µS/cm)	Total Cl^- (mg/L)
166	1382	14.9	1731.830	264.307
176	986	14.9	1235.589	96.368
190	821	15.1	1023.691	24.658
197	786	14.7	989.924	13.230
204	780	14.8	979.899	9.838
211	768	14.9	962.406	3.918
218	789	15.4	976.485	8.682
223	769	15.1	958.853	2.715
233	764	15.1	952.618	0.605

APPENDIX B: MATLAB Script

B-1. MATLAB Script for RMSE of all combinations in specific case.

```
clear all
[data,~,~]=xlsread('C:\Users\kaiyi\Desktop\Data.xlsx')
T=(data(3:12,1))'%Read observed time into vector T [min].
C=(data(3:12,2)/1000)' %Read observed conservation into vector C [mg/ml].
N=10 %10 samples other samples' concentration is 0 which will make the result as NaN.
x=30 %Observed point longitudinal location [cm].
M=557.7 %Total tracer mass [mg].
A=283.53 %Cross-section area [cm^2].
count=0
set=0
for i=1:1:N-2%Assume i<j<k<=N.
   t1=T(i)%Formula listed in Methodology.
   c1=C(i)
   for j=i+1:1:N-1
      t2=T(j)
      c2=C(j)
      for k=j+1:1:N
         t3=T(k)
         c3=C(k)
         w1=t2-t1
         w2=t3-t2
         r1=x*x*(1/t2-1/t1)
         r2=x*x*(1/t3-1/t2)
         b1=log(c1/c2*sqrt(t1/t2))
         b2=log(c2/c3*sqrt(t2/t3))
          if (w1*b2-w2*b1)<0
             break
          else
            if (r1*b2-r2*b1)>0
               break
            else
               Ue=sqrt(-4*(w1*b2-w2*b1)*(r1*b2-r2*b1))/2/(w1*b2-w2*b1)
               if ((w1*Ue*Ue+r1)/4/b1)<0
                  break
               else
                  count=count+1
                  DLe=(w1*Ue*Ue+r1)/4/b1
                  ne=M/2/A/c1/sqrt(DLe*pi*t1)*exp(-(x-Ue*t1)*(x-Ue*t1)/4/DLe/t1)
                  for z=1:1:N
                     Ce(z)=M/2/A/ne/sqrt(DLe*pi*T(z))*exp(-(x-Ue*T(z))*(x-
Ue*T(z))/4/DLe/T(z))%Substitute parameters for concentration estimation.
                     diff(z)=(C(z)-Ce(z))*(C(z)-Ce(z))
                  end
                  rmse=sqrt(sum(diff)/N)
                  RMSE(count,1:4)=[i j k rmse]
               end
            end
          end
      end
   end
   end
```

B-2. MATLAB Script for plotting estimated curve and observed data.

```
clear all
[data,~,~]=xlsread('C:\Users\kaiyi\Desktop\Data.xlsx')
T=(data(3:12,1))'
C=(data(3:12,2)/1000)'
N=10
x=30
M=557.7
A=283.53
% Based on MATLAB script in Appendix B-1, we already conclude the combination of i, j, and k for
the smallest RMSE.
i=1
   t1=T(i)
   c1=C(i)
j=4
   t2=T(j)
   c2=C(j)
k=5
   t3=T(k)
   c3=C(k)
   w1=t2-t1
   w2=t3-t2
   r1=x*x*(1/t2-1/t1)
   r2=x*x*(1/t3-1/t2)
   b1=log(c1/c2*sqrt(t1/t2))
   b2=log(c2/c3*sqrt(t2/t3))
   Ue=sqrt(-4*(w1*b2-w2*b1)*(r1*b2-r2*b1))/2/(w1*b2-w2*b1)
   DLe=(w1*Ue*Ue+r1)/4/b1
   ne=M/2/A/c1/sqrt(DLe*pi*t1)*exp(-(x-Ue*t1)*(x-Ue*t1)/4/DLe/t1)
   for z=1:1:N
      Ce(z)=M/2/A/ne/sqrt(DLe*pi*T(z))*exp(-(x-Ue*T(z))*(x-Ue*T(z))/4/DLe/T(z))
      diff(z)=(C(z)-Ce(z))*(C(z)-Ce(z))
   end
   rmse=sqrt(sum(diff)/N)
   for z=1:1:120
      t(z)=z
      Ce(z)=M/2/A/ne/sqrt(DLe*pi*t(z))*exp(-(x-Ue*t(z))*(x-Ue*t(z))/4/DLe/t(z))
   end
     sz=45
     MarkerSize=2
     hold on
     scatter(T,C,sz,'o','r');
     plot(t,Ce,'LineWidth',2)
     title('observed point x1=30cm u=0.25cm/cm')
     xlabel('Time(min)')
     ylabel('Concentration(mg/ml)')
```

B-3. MATLAB Script for plotting and RMSE calculation for two observed points.

```
clear all;
[data,~,~]=xlsread('C:\Users\kaiyi\Desktop\Data.xlsx')
T=(data(3:12,1))'
C=(data(3:12,2)/1000)'
T2=(data(2:14,3))'%Read observed time at x=70cm into Vector T2[min].
C2=(data(2:14,4)/1000)'%Read observed concentration at x=70cm into Vector C2[mg/ml].
N=10
N2=13%13 sets of observed data at x=70cm.
x=30
x2=70
M=557.7
A=283.53
i=1
   t1=T(i)
   c1=C(i)
j=4
   t2=T(j)
   c2=C(j)
k=5
   t3=T(k)
   c3=C(k)
   w1=t2-t1
   w2=t3-t2
   r1=x*x*(1/t2-1/t1)
   r2=x*x*(1/t3-1/t2)
   b1=log(c1/c2*sqrt(t1/t2))
   b2=log(c2/c3*sqrt(t2/t3))
   Ue=sqrt(-4*(w1*b2-w2*b1)*(r1*b2-r2*b1))/2/(w1*b2-w2*b1)
   DLe=(w1*Ue*Ue+r1)/4/b1
   ne=M/2/A/c1/sqrt(DLe*pi*t1)*exp(-(x-Ue*t1)*(x-Ue*t1)/4/DLe/t1)
   for z=1:1:N
      Ce(z)=M/2/A/ne/sqrt(DLe*pi*T(z))*exp(-(x-Ue*T(z))*(x-Ue*T(z))/4/DLe/T(z))
      diff(z)=(C(z)-Ce(z))*(C(z)-Ce(z))
   end
   rmse=sqrt(sum(diff)/N)
   for z=1:1:N2
      Ce2(z)=M/2/A/ne/sqrt(DLe*pi*T2(z))*exp(-(x2-Ue*T2(z))*(x2-Ue*T2(z))/4/DLe/T2(z))
      diff2(z)=(C2(z)-Ce2(z))*(C2(z)-Ce2(z))
   end
   rmse2=sqrt((sum(diff2))/N2)
   rmsetotal=sqrt((sum(diff)+sum(diff2))/(N2+N))
     for z=1:1:T(N)
        t(z)=z
        Ce(z)=M/2/A/ne/sqrt(DLe*pi*t(z))*exp(-(x-Ue*t(z))*(x-Ue*t(z))/4/DLe/t(z))
     end
     for z=1:1:T2(N2)
        t2(z)=z
        C2e(z)=M/2/A/ne/sqrt(DLe*pi*t2(z))*exp(-(x2-Ue*t2(z))*(x2-Ue*t2(z))/4/DLe/t2(z))
     end
     sz=45
     MarkerSize=2
     hold on
     scatter(T,C,sz,'o','r');
     scatter(T2,C2,sz,'o','k');
     plot(t,Ce,'-','LineWidth',2)
```

```
plot(t2,C2e,'LineWidth',2)
title('observed point x1=30cm u=0.25cm/cm')
xlabel('Time(min)')
ylabel('Concentration(mg/ml)')
```

B-4. MATLAB Script for plotting and RMSE calculation by curve-fitting method.

```
clear all
[data,~,~]=xlsread('C:\Users\kaiyi\Desktop\Data.xlsx')
T1=(data(3:12,1))'
C1=(data(3:12,2)/1000)'
N1=10
T2=(data(2:14,3))'
C2=(data(2:14,4)/1000)'
N2=13
x1=30;
x2=70;
M=557.7
A=283.53
n=0.13
for i=1:81 % for u from 0.4 cm/min to 0.8cm/min
   u=0.40+(i-1)*0.005

   for j=1:86  % for D from 0.05 cm^2/min to 0.22cm^2/min
      D=0.05+(j-1)*0.002

      for z=1:1:N1
         C1e(z)=M/2/A/n/sqrt(D*pi*T1(z))*exp(-(x1-u*T1(z))*(x1-u*T1(z))/4/D/T1(z))
         diff1(z)=(C1(z)-C1e(z))*(C1(z)-C1e(z))
      end
      for z=1:1:N2
         C2e(z)=M/2/A/n/sqrt(D*pi*T2(z))*exp(-(x2-u*T2(z))*(x2-u*T2(z))/4/D/T2(z))
         diff2(z)=(C2(z)-C2e(z))*(C2(z)-C2e(z))
      end
      rmse=sqrt((sum(diff1)+sum(diff2))/(N1+N2))

      RMSE(j,i)=rmse

   end
end

% Plot
clear all
[data,~,~]=xlsread('C:\Users\kaiyi\Desktop\Data.xlsx')
T1=(data(3:12,1))'
C1=(data(3:12,2)/1000)'
N1=10
T2=(data(2:14,3))'
C2=(data(2:14,4)/1000)'
N2=13
x1=30;
x2=70;
M=557.7
A=283.53
n=0.13
i=36 % Result from workspace concluded by the script above
   u=0.40+(i-1)*0.005

j=4
      D=0.1+(j-1)*0.002

      for z=1:1:N1
```

```
    C1e(z)=M/2/A/n/sqrt(D*pi*T1(z))*exp(-(x1-u*T1(z))*(x1-u*T1(z))/4/D/T1(z))
    diff1(z)=(C1(z)-C1e(z))*(C1(z)-C1e(z))
 end
 for z=1:1:N2
    C2e(z)=M/2/A/n/sqrt(D*pi*T2(z))*exp(-(x2-u*T2(z))*(x2-u*T2(z))/4/D/T2(z))
    diff2(z)=(C2(z)-C2e(z))*(C2(z)-C2e(z))
 end
 rmse1=sqrt((sum(diff1))/(N1))
 rmse2=sqrt((sum(diff2))/(N2))
 RMSE=sqrt((sum(diff1)+sum(diff2))/(N1+N2))

for z=1:1:200
   t2(z)=z
   C2ee(z)=M/2/A/n/sqrt(D*pi*t2(z))*exp(-(x2-u*t2(z))*(x2-u*t2(z))/4/D/t2(z))
end
for z=1:1:120
   t1(z)=z
   C1ee(z)=M/2/A/n/sqrt(D*pi*t1(z))*exp(-(x1-u*t1(z))*(x1-u*t1(z))/4/D/t1(z))
end
sz=45
MarkerSize=2
hold on
scatter(T1,C1,sz,'o','r');
scatter(T2,C2,sz,'o','k');
plot(t1,C1ee,'-*','LineWidth',1)
plot(t2,C2ee,'LineWidth',2)
xlabel('Time(min)')
ylabel('Concentration(mg/ml)')
```